(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 October 2003 (30.10.2003)

PCT

(10) International Publication Number
WO 03/089649 A1

(51) International Patent Classification[7]: C12N 15/85, C07K 16/18, C12N 15/62, 5/10, A61K 39/00

(21) International Application Number: PCT/GB03/01659

(22) International Filing Date: 17 April 2003 (17.04.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0209023.1 19 April 2002 (19.04.2002) GB

(71) Applicant *(for all designated States except US)*: **OXFORD BIOMEDICA (UK) LIMITED** [GB/GB]; Medawar Centre, Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **KINGSMAN, Susan** [GB/GB]; Oxford Biomedica (UK) Limited, Medawar Centre, Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **CARROLL, Miles** [GB/GB]; Oxford Biomedica (UK) Limited, Medawar Centre, Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **MYERS, Kevin** [GB/GB]; Oxford Biomedica (UK) Limited, Medawar Centre, Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **DRURY, Noel** [GB/GB]; Oxford Biomedica (UK) Limited, Medawar Centre, Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB).

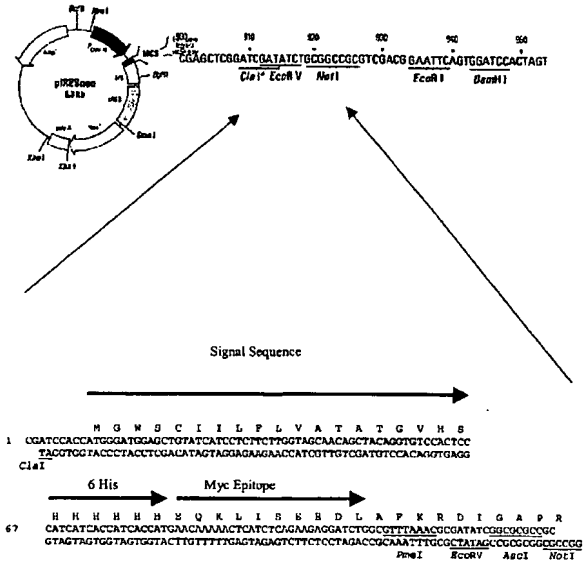
(74) Agents: **MASCHIO, Antonio** et al.; D Young & Co, 21 New Fetter Lane, London EC4A 1DA (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,

[Continued on next page]

(54) Title: EXPRESSION VECTOR COMPRISING A SIGNAL SEQUENCE AND AN AMINO-TERMINAL PEPTIDE TAG



(57) Abstract: The invention relates to an expression vector comprising an amino-terminal tag sequence and a signal sequence operably linked to a nucleotide sequence of interest, wherein the amino-terminal tag sequence is inserted between the signal sequence and the nucleotide sequence of interest.

WO 03/089649 A1

SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv)) for US only*

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

EXPRESSION VECTOR COMPRISING A SIGNAL SEQUENCE AND AN AMINO-TERMINAL PEPTIDE TAG

FIELD OF INVENTION

The present invention relates to an expression vector, methods of use and products obtained therefrom.

In particular, the present invention relates to an expression vector in which an amino-terminal tag sequence is inserted between a signal sequence and a nucleotide sequence of interest.

BACKGROUND TO THE INVENTION

A promising new development in cancer treatment is antibody-targeted chemotherapy (ATC), in which a cytotoxic agent is linked to an antibody that recognises distinctive molecules called tumour associated antigens (TAAs) expressed by tumour cells e.g. leukaemic cells (Appelbaum (1999) *Semin. Hematol.* 36, 2-B). Such an approach allows the specific delivery of a cytotoxic drug to the tumour cells and is likely to result in fewer side effects.

A variety of antigens have been identified for use in ATC. A critical factor in the success of this treatment strategy is careful selection of the tumour antigen for targeting. Ideally, the targeted antigen should be differentially expressed on tumour cells and have a lower level of expression on normal cells. In practice, targeting of a subset of normal cells is largely unavoidable. However, in order to allow the necessary renewal of the normal cell population, the stem cell precursor of the lineage must not share the antigen that is targeted on the tumour cells. Another important factor is the stability of the antigen on the cell surface. The ideal target should not shed from the cell surface, or secrete into the circulation.

Monoclonal antibodies (mAbs) against antigens expressed on tumour cell surfaces have been used to deliver chemotherapeutic drugs (Dubowchik & Walker (1999)

Pharmacol. Ther. 83, 67-123), potent plant and bacterial toxins (Brinkmann (2000) *In Vivo* 14, 21-27) and radionuclides (Illidge & Johnson (2000) *Br. J. Haematol.* 108, 679-688) to tumours. A major advance in mAb based targeting strategies was made in developing Mylotarg. Mylotarg is a highly specific antibody that recognises a cell-surface molecule called "CD33" which is abundant on acute myeloid leukemia (AML) cells but absent from normal blood stem cells, the seeds from which normal blood and immune cells originate.

AML is a virulent and often fatal form of cancer in which certain white blood cells become cancerous and rapidly accumulate in the bone marrow, preventing normal marrow from growing and functioning properly. It is among the most serious forms of adult leukemia, with a relatively high fatality rate. Most patients with AML require intensive standard chemotherapy to achieve remission, and some also must undergo bone marrow transplants. Even after such intensive treatment, up to half of all AML patients have residual leukemic cells or experience a relapse. Because standard chemotherapy drugs to treat AML are non-specific (i.e. normal as well as malignant cells are destroyed) patients who receive the therapy tend to become very sick.

The CD33 antigen is an appropriate target because AML blast cells express the CD33 antigen in more than 90 % of patients, whereas hematopoietic stem cells, lymphoid cells, and nonhematopoietic cells do not express the CD33 antigen (Dinndorf *et al.* (1986) *Blood* 67, 1048-1053; Griffin *et al.* (1984) *Leuk. Res.* 6, 521-534; Pelper *et al.* (1988) *Blood* 72, 314-321). The CD33 antigen, a member of the immunoglobulin superfamily, is a membrane-bound glycoprotein. The humanised murine CD33 antibody is linked to a novel and extremely potent chemotherapy agent known as "calicheamicin". The antibody selectively targets leukemic blast cells and delivers calicheamicin to them. As a result, the leukemic cells are destroyed but the cells that are responsible for replenishing normal blood cells are spared. Vincent *et al.* (2001) *Blood* 97, 3197-3203 have demonstrated that Mylotarg rapidly binds to CD33 antigenic sites expressed on leukemic and normal myeloid cells in blood, but not to lymphocytes.

Other promising agents that have been through Phase III clinical trials for the treatment of B-cell lymphomas include Zevalin (Wiseman *et al.* (2000) *Eur. J. Nucl. Med.* 27, 766-777) and Bexxar (Hainsworth (2000) *Oncologist* 5 376-384). The unconjugated mAb Herceptin has demonstrated significant levels of activity for the treatment of metastatic breast carcinoma (Stebbing *et al.* (1998) *Cancer Treat. Rev.* 26, 287-290) and Panorex for use in patients with colorectal cancer which has spread to the nearby lymph glands (Welt & Ritter (1999) *Semin. Oncol.* 26 683-690).

In recent years many mAbs immunologically reactive with TAAs have been isolated (Allum *et al.* (1986) *Surg. Ann.* 18: 41064; Houghton *et al.* (1986) *Semin. Oncol.* 13: 165-179). By way of example, US 5597707 teaches an isolated and substantially purified glycoprotein tumour-associated antigen. The antigen is specifically recognised by the murine mAb L6.

The antigens bound by these mAbs and characterisation of the binding pattern of the mAbs in most cases are incomplete and somewhat uncertain. Some of the more completely characterised mAbs have been found to bind to a variety of TAAs including glycoproteins, glycolipids and mucins (Fink *et al.* (1984) *Prog. Clin. Pathol.* 9: 121-133).

ATC is therefore a potential treatment against cancer, in which an important step is the identification of TAAs that are expressed by tumours for the specific delivery of cytotoxic drugs. It is important to be able to determine whether a TAA will be of use in cancer immunotherapy. Furthermore, there is a bottle neck in the screening of putative TAAs in that a specific mAb is required for each TAA that is identified.

SUMMARY OF THE INVENTION

The present invention is based upon the surprising finding that the expression profile of novel cDNAs - such as TAAs - may be determined to establish whether the cDNA is a disease target for use in cancer immunotherapy. Moreover, the potential utility of

novel cDNAs as disease targets may be assessed without the requirement of specific mAbs to each disease target.

Accordingly, the present invention may be used to establish the cellular location of the novel cDNA gene product, whether membrane bound proteins are internalised on binding an antibody, whether the protein is secreted, the size of the gene product and optionally the detection of cytotoxic effects of an antibody binding to a cell expressing a tagged target antigen.

SUMMARY ASPECTS OF THE PRESENT INVENTION

In a first aspect, there is provided an expression vector comprising an amino-terminal tag sequence and a signal sequence operably linked to a nucleotide sequence of interest, wherein the amino-terminal tag sequence is inserted between the signal sequence and the nucleotide sequence of interest.

In a second aspect, there is provided a host cell comprising the expression vector of the present invention.

In a third aspect, there is provided a method for determining the utility of a nucleotide sequence of interest as a disease target, comprising the steps of: (a) providing an expression vector according to the present invention, comprising a nucleotide sequence of interest; (b) transfecting a host cell with the expression vector; (c) selecting for host cells expressing the nucleotide sequence of interest; and (d) determining the expression profile of the protein expressed from the nucleotide sequence of interest.

In a fourth aspect, there is provided a method for determining the utility of a nucleotide sequence of interest as a disease target, comprising the steps of: (a) providing an expression vector according to the present invention comprising a nucleotide sequence of interest; (b) transfecting a host cell with the expression vector; (c) selecting for host cells expressing the nucleotide sequence of interest; and (d) determining the expression profile of the protein expressed from the nucleotide

sequence of interest; and (e) purifying the protein expressed from the nucleotide sequence of interest

In a fifth aspect, there is provided a method for preparing an antibody against a disease target polypeptide identified by the methods of the present invention comprising the step of expressing the nucleotide sequence encoding the disease target polypeptide.

In a sixth aspect, there is provided an antibody prepared by the methods of the present invention.

In a seventh aspect, there is provided a kit for determining the utility of a nucleotide sequence of interest as a disease target comprising an expression vector according to the present invention.

In an eighth aspect, there is provided a library of nucleotide sequences of interest in an expression vector according to the present invention.

In a ninth aspect, there is provided a method of preparing a library of nucleotide sequences of interest comprising the steps of: (a) treating a nucleic acid sample to fragment the nucleic acid contained therein; (b) ligating the nucleic acid fragments formed in step (a) in to a vector; (c) transforming the vector into a host cell; (d) selecting for host cells comprising the vector; (e) identifying the signal sequence; (f) amplifying the gene immediately downstream from the signal sequence; (g) sub-cloning the gene in to an expression vector according to the present invention; and (h) determining the expression profile of the protein expressed from the nucleotide sequence of interest.

In a tenth aspect, there is provided a disease target identified by the methods of the present invention.

In an eleventh aspect, there is provided an antibody linked to an agent that binds to the disease target identified by the methods of the present invention.

In a twelfth aspect, there is provided an antibody according to the present invention for use in the treatment of disease.

In a thirteenth aspect, there is provided a use of an antibody according to the present invention in the manufacture of a composition for the treatment of a disease.

In a fourteenth aspect, there is provided a method of treating a disease in a subject comprising administering to said subject a medically effective amount of an antibody according to the present invention.

Other aspects of the present invention are presented in the accompanying claims and in the following description and discussion. These aspects are presented under separate section headings. However, it is to be understood that the teachings under each section heading are not necessarily limited to that particular section heading.

PREFERRED EMBODIMENTS

Preferably, the nucleotide sequence of interest is a disease target. More preferably, the nucleotide sequence of interest is a tumour associated antigen.

Preferably, the amino-terminal tag sequence comprises a 6-His sequence.

Preferably, the amino-terminal tag sequence comprises a c-Myc epitope.

Preferably, the amino-terminal tag sequence comprises the sequence as set forth in Seq ID No.3 or a fragment, mutant, variant or homologue thereof.

Preferably, the signal sequence comprises the sequence as set forth in Seq ID No.5 or a fragment, mutant, variant or homologue thereof.

Preferably, the expression vector comprises a selectable marker.

Preferably, the expression vector is pIRES_STAR comprising the sequence as set forth in Seq ID No.2 or a fragment, mutant, variant or homologue thereof.

Preferably, step (d) of the method according to the third or fourth aspects of the present invention comprises FACS analysis of the binding of a labelled anti-amino-terminal tag antibody.

Preferably, step (d) of the method according to the third or fourth aspects of the present invention comprises microscopic analysis of a labelled anti-amino-terminal tag antibody.

Preferably, the antibody is a monclonal antibody.

Preferably, antibody is a humanised antibody.

Preferably, the antibody is linked to an agent.

Preferably, the disease is cancer.

DESCRIPTION OF THE DRAWINGS

Figure 1 is a diagrammatic representation of the construction of the expression vector pIRES_STAR. A double stranded oligonucleotide coding for the murine Igκ leader sequence followed by 6 histidines, a myc epitope and several 8 base recognition restriction enzyme sites is inserted into pIRESneo. The oligonucleotide is made from two complementary single stranded oligos that are annealed to give overhanging ends that are compatible with the *Cla*I and *Not*I sites of pIRESneo.

Figure 2 is a diagrammatic representation of the amplification and cloning of h5T4. Primers are designed to amplify full length human 5T4 without its signal sequence (S-FL) and the 5T4 gene without its transmembrane domain and cytoplasmic tail as well as minus the signal sequence (S-Tm-).

Figure 3 is a series of photographs of an immunostain of Chinese Hamster Ovary (CHO) cells transfected with full length and Tm-h5T4 at 10X magnification. In more detail, a)-d) are 48h post-transfection (transient) stains. a) CHO cells transfected with pIRES_h5T4EKMH stained with anti-Myc (1:200), b) same stained with H8 (hybridoma supernatant 1:20), c) CHO cells transfected with pIRES_STARh5T4 stained with anti-Myc, d) same stained with H8 e) CHO_STARh5T4 stable transfectant (clone 4) stained with anti-Myc f) CHO_h5T4EKMH stained with anti-Myc. Secondary antibody of rabbit-anti-mouse-HRP was used (1:1000)

Figure 4 is a series of photographs of an immunocytochemistry analysis of CHO_STARh5T4 and CHO_h5T4EKMH. In more detail, the photographs show the immunocytochemistry analysis of CHO_STARh5T4 (clone 4) cells with H8 anti-h5T4 antibody (1:1000), anti-Myc (1:200) or anti-pentaHis (1:200) and secondary antibody of rabbit-anti-mouse-FITC (1:20) (40X magnification). a) CHO_h5T4EKMH/H8 b) CHO_STARh5T4 (clone 4)/H8 c) CHO_h5T4EKMH/anti-Myc d) CHO_STARh5T4 (clone 4)/ anti-Myc e) CHO_h5T4EKMH/anti-pentaHis f) CHO_STARh5T4 (clone 4)/ anti-pentaHis.

Figure 5A is a series of diagrams representing the FACS analysis of CHO_STARh5T4 cell lines. a) Fluorescence histogram of CHO_h5T4EKMH with H8 (1:200) or anti-Myc (1:500); b) Fluorescence histogram of CHO_h5T4EKMH with H8 (1:200) or anti-pentaHis (1:500); c) Fluorescence histogram of CHO_STARh5T4 with H8 (1:200) or anti-Myc (1:500); d) Fluorescence histogram of CHO_STARh5T4 with H8 (1:200) or anti-pentaHis (1:500).

Figure 5B is a series of diagrams representing the FACS analysis of h5T4 internalisation by CHO_STARh5T4 cell lines in response to H8 (1:1000) and anti-pentaHis (1:100). a) CT26_h5T4 cells with H8; b) CHO_h5T4EKMH cells with H8; c) CHO_h5T4EKMH cells with anti-pentaHis; d) CHO_STARh5T4 cells with H8; e) CHO_STARh5T4 cells with anti-pentaHis.

Figure 6 is a series of photographs of the Western analysis of supernatants or cell lysates from CHO_STARh5T4 cells.

a) Supernatants; anti-Myc 1:1000, rabbit anti-mouse-HRP 1:10,000, b) Supernatants; H8 anti-h5T4 1:1000, rabbit anti-mouse-HRP 1:10,000. Lane 1 CHO_STARh5T4 (clone 4) unreduced; Lane 2 CHO_STARh5T4 (clone 4) reduced; Lane 3 CHO_STARh5T4TM- (clone 9) unreduced; Lane 4 CHO_STARh5T4TM- (clone 9) reduced; Lane 5 CHO unreduced; Lane 6 CHO reduced.

c) Cell lysates; anti-Myc d) Cell lysates; H8. Numbers on the left of the blots indicate molecular weight markers. Lane A, CHO unreduced; Lane B, CHO reduced; Lane C, CHO_h5T4EKMH unreduced; Lane D, CHO_h5T4EKMH reduced; Lane E, CHO_STARh5T4 (clone 4) unreduced; Lane F, CHO_STARh5T4 (clone 4) reduced.

DETAILED DESCRIPTION OF THE INVENTION

EXPRESSION VECTOR

As used herein, the term "expression vector" means a vector capable of *in vivo* or *in vitro* expression of a nucleotide sequence. For example, the nucleotide sequence may be a nucleotide sequence of interest that is present in the expression vector of the present invention. Nucleotide sequences are linked to a regulatory sequence, which is capable of providing for the expression of the coding sequence. By way of example, the present invention covers a vector comprising a nucleotide sequence of interest operably linked to such a regulatory sequence.

The term "regulatory sequences" includes promoters and enhancers and other expression regulation signals.

The expression vector may be transformed into a suitable host cell as described below to provide for expression of a polypeptide.

Preferably, the expression vector is a plasmid provided with an origin of replication, a promoter for the expression of a nucleotide sequence and optionally a regulator of the promoter.

In a preferred embodiment, the expression vector of the present invention comprises a selectable marker in order to select cells comprising the expression vector. Many different selectable markers may be used and include, but are not limited to, the bacterial neomycin and hygromycin phosphotransferase genes which confer resistance to G418 and hygromycin respectively; a mutant mouse dihydrofolate reductase gene which confers resistance to methotrexate; the bacterial *gpt* gene which allows cells to grow in medium containing mycophenolic acid, xanthine and aminopterin; the bacterial *hisD* gene which allows cells to grow in medium without histidine but containing histidinol; the multidrug resistance gene (*mdr*) which confers resistance to a variety of drugs; and the bacterial genes which confer resistance to puromycin or phleomycin. Preferably, the selectable marker is a gene encoding the neomycin phosphotransferase gene.

A nucleotide sequence may be incorporated into the expression vector. By way of example, the nucleotide sequence may be a nucleotide sequence of interest incorporated into the expression vector. This may be facilitated by the inclusion of restriction enzyme sites. In a preferred embodiment, the expression vector comprises the restriction enzyme sites *Pme*I, *Eco*RV, *Asc*I and *Not*I by incorporation of the sequence Seq. ID No.1 or a fragment, mutant, variant or homologue thereof:

Seq ID No. 1

5' gcgtttaaacgcgatatcggcgcgccgc 3'

Preferably, PCR amplified cDNAs are directionally cloned using the blunt *Pme*I site at the 5' end and the overhanging *Asc*I at the 3' end. Thus, the 5' primer will not require an additional restriction site as Pfu polymerase leaves a blunt ended product, whilst the 3' primer will incorporate an additional *Asc*I site. Other restriction sites may be

included in the expression vector using methods that are well known in the art to allow alternative cloning strategies if required.

Expression vectors may be used to replicate and express a nucleotide sequence. Expression may be controlled using control sequences, which include promoters/enhancers and other expression regulation signals. Prokaryotic promoters and promoters functional in eukaryotic cells may be used. Tissue-specific or stimuli-specific promoters may be used. Chimeric promoters may also be used comprising sequence elements from two or more different promoters.

The term "promoter" is used in the normal sense of the art.

Suitable promoting sequences are preferably strong promoters including those derived from the genomes of viruses - such as polyoma virus, adenovirus, fowlpox virus, bovine papilloma virus, avian sarcoma virus, cytomegalovirus (CMV), retrovirus and Simian Virus 40 (SV40) – or from heterologous mammalian promoters - such as the actin promoter or ribosomal protein promoter. Transcription of a gene may be increased further by inserting an enhancer sequence into the vector. Enhancers are relatively orientation- and position-independent, however, one may employ an enhancer from a eukaryotic cell virus – such as the SV40 enhancer on the late side of the replication origin (bp 100-270) and the CMV early promoter enhancer. The enhancer may be spliced into the vector at a position 5' or 3' to the promoter, but is preferably located at a site 5' from the promoter.

Preferably, the promoter is a CMV major immediate early promoter/enhancer.

Hybrid promoters may also be used to improve inducible regulation of the expression vector.

The promoter can additionally include features to ensure or to increase expression in a suitable host. For example, the features can be conserved regions e.g. a Pribnow Box or a TATA box. The promoter may even contain other sequences to affect (such as to

maintain, enhance, decrease) the levels of expression of a nucleotide sequence. Suitable other sequences include the Sh1-intron or an ADH intron. Other sequences include inducible elements - such as temperature, chemical, light or stress inducible elements. Also, suitable elements to enhance transcription or translation may be present.

The expression vector of the present invention comprises a signal sequence and an amino-terminal tag sequence operably linked to a nucleotide sequence of interest.

As used herein, the term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. For example, a regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

In a preferred embodiment, the expression vector is pIRES_STAR (as illustrated in Figure 1) wherein STAR refers to Surface Tumour Antigen Rapid-evaluation technology. pIRES_STAR is constructed by inserting into pIRESneo (Genbank Accession Number U89673), a double stranded oligonucleotide comprising the sequence Seq ID No. 2 or a fragment, mutant, variant, or homologue thereof:

Seq ID No. 2

5' atcgatccaccatgggatggagctgtatcatcctcttcttggtagcaacagctacaggtgtccactcccatcatcaccat caccatgaacaaaaactcatctcagaagaggatctggcgtttaaacgcgatatcggcgcgccgcggccgc 3'

AMINO-TERMINAL TAG

The term "amino-terminal tag" refers to a nucleotide sequence comprising a detection means or a purification means. Preferably, the amino-terminal tag comprises both a purification means and a detection means. In some instances, the detection means may also be used as a purification means and *vice versa*.

Normally, surface-expressed proteins are tagged at the carboxy-terminus due to complications of using the amino-terminus, which involves removal of signal sequences, and possible structural effects that the tag may have on the extra-cellular domain of the protein. Additionally, in the case of complex glycoproteins it is quite likely that an amino-terminal located tag will not be accessible to the anti-tag antibody, which may prevent the subsequent analysis of the target protein.

In one aspect of the present invention, the amino-terminal tag sequence is inserted between the signal sequence and the nucleotide sequence of interest, placing the amino-terminal tag upstream of the nucleotide sequence of interest and downstream of the signal sequence. In other words, the amino-terminal tag of the present invention is inserted between the leader sequence and the predicted N-terminus of the mature, processed protein, since membrane bound and secreted proteins have a leader sequence to direct them to the cell membrane which is subsequently cleaved off.

The term "detection means" refers to any tag that can be used for the labelling and detection of expressed recombinant proteins in bacteria, yeast, insect and mammalian cell systems and includes 6-His tagged proteins, glutathione S-transferase (GST) tagged proteins and epitope tagged proteins.

Preferably, epitope tags are used which due to their small size, are unlikely to affect the tagged protein's biochemical properties. More preferably, the c-Myc epitope tag is used.

c-Myc-tagged proteins may be detected using a c-Myc-Tag polyclonal antibody. Polyclonal antibodies may be produced by immunising rabbits with a synthetic peptide (KLH coupled) corresponding to residues 410–419 of human c-Myc (EQKLISEEDL) which recognise the Myc-tag fused to the C-terminus and N-terminus of targeted proteins in transfected mammalian cells. Antibodies may be purified by protein A and peptide affinity chromatography. c-Myc-tagged proteins may also be detected using monoclonal antibodies.

Preferably, the monoclonal antibody is the 9E10 monoclonal antibody, which is commercially available. This antibody is highly specific and unlikely to alter the reactivity of the tagged protein.

The antibodies may be detected using various methods known in the art - such as Western blotting, immunoprecipitation, immunofluorescence microscopy and ELISA.

The term "purification means" refers to any "tag" that can be used for the purification of expressed recombinant proteins in bacteria, yeast, insect and mammalian cell systems and include, but are not limited to, 6-His tagged proteins and glutathione S-transferase (GST) tagged proteins.

For example, the QIAexpress protein purification and expression system (QIAGEN) uses nickel nitrilotriacetic acid (Ni-NTA) to bind proteins with a 6-His tag. The tag and the purification procedure are compatible with any expression system – such as bacteria, mammalian cell, baculovirus, or yeast. Because the histidine-nickel interaction does not depend upon the structure of the tagged protein, the system can be used under denaturing conditions, often simplifying purification. The tag's affinity for nickel ions immobilised on the resin means that, even if the protein expression level is low, the protein can be purified to homogeneity under native or under denaturing conditions.

Preferably, the tag used for purification is a 6-His tag.

By way of example, a 6-His tagged protein may be purified using the following method. Cells are spun down and washed in PBS. The cell pellet is resuspended in lysis buffer (50 mM Tris pH 8.0, 1 % NP-40, 1 mM PMSF, 1 µg/ml leupeptin) and sonicated. Lysate are again spun and the supernatant is bound to 0.5 ml Ni-NTA pre-equilibrated with lysis buffer and mixed at 4°C for 1 hour to overnight. The slurry is transferred to a column and the flow-through is collected. The column is washed with Wash Buffer 1 (20 ml 50 mM Tris pH 8.0, 500 mM NaCl and 1 mM PMSF). Optionally, the column is also washed with Wash Buffer 2 (20 ml 50 mM Tris pH 8.0,

500 mM NaCl, 5 mM imidazole 1 mM PMSF). The protein of interest is eluted with Elution Buffer (5 ml 50 mM Tris pH 8.0, 500 mM NaCl, 1 M imidazole, 1 mM PMSF) and fractions are collected.

In a preferred embodiment, the amino-terminal tag comprises 6 histidines and a c-Myc epitope and comprises the sequence Seq ID No. 3 or a fragment, mutant, variant, or homologue thereof:

Seq ID No. 3:

5'catcatcaccatcaccatgaacaaaaactcatctcagaagaggatctg 3'

SIGNAL SEQUENCE

As used herein, the term "signal sequence" is synonymous with the terms "signal peptide" and "leader sequence".

Protein transport is an essential process for both prokaryotic and eukaryotic cells. Transport of an individual protein usually occurs via an amino-terminal signal sequence, which directs, or targets, the protein from its ribosomal assembly site to a particular cellular or extracellular location. Transport may involve any combination of several of the following steps: contact with a chaperone, unfolding, interaction with a receptor and/or a pore complex, addition of energy, and refolding. Moreover, an extracellular protein may be produced as an inactive precursor. Once the precursor has been exported, removal of the signal sequence by a signal peptidase activates the protein.

Some examples of the protein families which are known to have signal sequences are receptors (nuclear, 4 transmembrane, G protein coupled, and tyrosine kinase), cytokines (chemokines), hormones (growth and differentiation factors), neuropeptides and vasomediators, protein kinases, phosphatases, phospholipases, phosphodiesterases, nucleotide cyclases, matrix molecules (adhesion, cadherin, extracellular matrix

molecules, integrin, and selectin), G proteins, ion channels (calcium, chloride, potassium, and sodium), proteases, transporter/pumps (amino acid, protein, sugar, metal and vitamin; calcium, phosphate, potassium, and sodium) and regulatory proteins.

It is possible in principle to distinguish between two different types of signal sequences: a "hydrophobic" type and a "hydrophilic" type. The "hydrophobic" group of signal sequences usually comprises about 13-30 amino acids, whereas the "hydrophilic" group comprises about 12-70 amino acids. The signal sequence of the "hydrophobic" type can be divided into three structural elements. It is composed of a relatively hydrophilic amino-terminus with one or two basic amino acids, of an apolar, mostly hydrophobic block of seven or eight amino acids, and of a relatively hydrophilic COOH-terminus which is terminated by an amino acid with a small side-chain. Such "hydrophobic" signal sequences guide proteins through the membrane of the endoplasmic reticulum (ER) and through bacterial membranes. Although bacterial and ER signal sequences differ slightly from one another, they are functionally interchangeable. The structure of the "hydrophilic" type differs greatly from that of the "hydrophobic" type: there are no lengthy uninterrupted sections of hydrophobic amino acids in the "hydrophilic" type, but there are usually many basic and hydroxylated amino acids and few or no acidic amino acids. The "hydrophilic" type of signal sequences guides proteins into mitochondria, chloroplasts and, possibly, into peroxisomes too.

Although the "hydrophobic" type of signal sequences of prokaryotic and eukaryotic origin have common characteristics and may be functionally interchangeable, there are also observable differences: thus, most of the prokaryotic signal sequences that are known have, by comparison with the "hydrophobic" type of eukaryotic signal sequences, a lower hydrophobicity in the apolar section plus, usually, an additional basic amino acid in the amino-region. This is possibly the reason why the natural signal sequence of a heterologous protein is usually less efficiently recognized and processed in microorganisms than is a bacterial signal sequence.

The signal sequence is often (but not always) removed from the mature protein once the sorting decision has been executed. Thus, according to the present invention, the signal sequence is inserted upstream from the amino-terminal tag which is inserted between the signal sequence and the predicted N-terminus of the mature, processed protein. If the tag were placed at the carboxy- terminus, and therefore intracellular, it would be impossible for the expression profile to be determined using an anti-tag antibody.

In accordance with the present invention, the signal sequence may be from any protein – such as a mammalian protein – that is available in a database, such as GenBank or other database. A person skilled in the art will appreciate that the signal sequence for a given protein may be identified using, for example, the Signal P program as described below.

The signal sequence may be a heterologous signal sequence *ie*. a signal sequence that does not occur naturally as part of the organism in which it is present. Typically, the heterologous signal sequence is not endogenous to the cell into which it is introduced, but is obtained from another cell.

Preferably, the signal sequence localises a protein expressed from the nucleotide sequence of interest to the surface of a cell. More preferably, the signal sequence localises a protein expressed from the nucleotide sequence of interest to the plasma membrane *ie*. the signal sequence is a plasma membrane signal sequence.

The signal sequence may be any of the HLA genes including HLA B*4409 (Accession No. AJ309937) comprising the sequence Seq. ID No. 4 or a fragment, mutant, variant, or homologue thereof may be used:

Seq ID No. 4

MRVTAPRTLLLLLWGAVALTETWA

Such a signal sequence may be inserted in to the expression vector of the present invention using routine methods known in the art.

Preferably, the signal sequence is the murine Igκ signal sequence comprising the amino acid sequence Seq. ID No. 5 or a fragment, mutant, variant, or homologue thereof:

Seq. ID No. 5

MGWSCIILFLVATATGVHS

According to the present invention, novel nucleotide sequences, for example, cDNAs are obtained in standard cloning vectors using methods known in the art.

The leader sequence may be identified using the Signal P program (http://www.cbs.dtu.dk/services/SignalP/). The method has been described in Nielsen *et al.* (1997) *Protein Engineering* 10, 1-6. The SignalP program predicts the presence and location of signal peptide cleavage sites in amino acid sequences from different organisms: Gram-positive prokaryotes, Gram-negative prokaryotes, and eukaryotes. The method incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks. Other methods for predicting the signal sequence are also known including the algorithm formulated by Chou (*Peptides* (2001) 22, 1973-9) which may also serve as a complementary tool to the existing algorithms for signal peptide prediction.

By way of example only, the Signal P program has been used to identify the leader sequence for the following proteins:

ACCESSION NP_000735; leader sequence = MELGGPGAPR LLPPLLLLLG TGLLRASS; ACCESSION NP_008820; leader sequence = MGLSNILFVM AFLLSG; ACCESSION NP_000555; leader sequence = MIIVAHVLLI LLGATEILQA; ACCESSION NP_004548; leader sequence = MQPPSLLLLL

LLLLLCVSVV RPR; ACCESSION NP_002812; leader sequence = MGAARGSPAR PRRLPLLSVL LLPLLGGTQT; ACCESSION NP_002380; leader sequence = MEPPGRRECP FPSWRFPGLL LAAMVLLLYS FSDA.

Following this analysis, the gene immediately downstream from the signal sequence may be amplified using various methods known in the art. By way of example, a specific PCR primer may be designed to amplify the gene immediately downstream from the signal sequence, which is used in conjunction with a generic primer located downstream of the cloning site in the vector. This will give a product representing the coding sequence of the protein product, incorporating its own termination codon, but minus the native signal sequence. The amplified product may then be cloned into the expression vector of the present invention for expression profile analysis.

NUCLEOTIDE SEQUENCE

The present invention involves the use of nucleotide sequences, which may be available in databases. These nucleotide sequences may be used to express amino acid sequences.

The nucleotide sequence may be DNA or RNA of genomic, synthetic or recombinant origin e.g. cDNA. The nucleotide sequence may be double-stranded or single-stranded whether representing the sense or antisense strand or combinations thereof.

The nucleotide sequence may be prepared by use of recombinant DNA techniques (*e.g.* recombinant DNA).

The nucleotide sequence may be the same as the naturally occurring form, or may be derived therefrom.

The nucleotide sequence may be a nucleotide sequence of interest *i.e.* a nucleotide sequence representing the coding sequence of the protein product, incorporating its own termination codon, but minus the native signal sequence.

Preferably the nucleotide sequence of interest is a disease target.

DISEASE TARGET

Aspects of the present invention relate to disease targets identified by the methods of the present invention. As used herein, the term "disease target" refers to any nucleotide sequence of interest that is associated with any anatomical abnormality or impairment of the normal function of an organism (e.g. a human) or any of its parts other than one arising directly from personal injury. The disease may be caused by environmental factors, infective agents, genetic disease or any combination thereof and may include cancer.

Preferably, the disease target encodes a protein (e.g. an antigen) that is associated with the surface of a cell, becomes internalised in the cell, is not secreted from the cell and is the same size before and after entering the cell. More preferably, the disease target is a tumour associated antigen (TAA) that is associated with the surface of a tumour cell, becomes internalised in the tumour cell, is not secreted from the tumour cell and is the same size before and after entering the tumour cell.

TAA

The term "tumour associated antigen (TAA)" is used herein to refer to any TAA or antigenic peptide thereof, the antigen being one that is expressed by the tumour itself or cells associated with the tumour. The term includes antigens that distinguish the tumour cells from their normal cellular counterparts where they may be present in trace amounts.

Thus, TAAs are surface molecules that are differentially expressed in tumour cells relative to non-tumor tissues. TAAs make tumour cells immunologically distinct from normal cells and provide diagnostic and therapeutic targets for human cancers. Several mAbs have been identified which react specifically with cancerous cells e.g.

T-cell acute lymphoblastic leukemia and neuroblastoma (Minegishi *et al.* (1989) *Leukemia Res.* 13: 43-51; Takagi *et al.* (1995) *Int. J. Cancer* 61: 706-715).

TAAs have been characterised either as membrane proteins or altered carbohydrate molecules of glycoproteins and glycolipids, however their functions remain largely unknown. One TAA family, the transmembrane 4 superfamily (TM4SF), usually has four well-conserved membrane-spanning regions, certain cysteine residues and short sequence motifs. There is evidence that TM4SF antigens exist in close association with other important membrane receptors including CD4 and CD8 of T cells (Imai & Yoshie (1993) *J. Immunol.* 151, 6470-6481). It has also been suggested that TM4SF antigens may play a role in signal transduction which in turn, affects cell development, activation and motility. Examples of TM4SF antigens include human melanoma-associated antigen ME491, human and mouse leukocyte surface antigen CD37, and human lymphoblastic leukemia-associated TALLA-1 (Hotta, H. et al. (1988) Cancer Res. 48, 2955-2962; Classon, B. J. et al. (1989) J. Exp. Med. 169: 1497-1502; Tomlinson, M. G. et al. (1996) Mol. Immun. 33: 867-872; Takagi, S. et al. (1995) Int. J. Cancer 61: 706-715).

Further examples of TAAs include but are not limited to MART-1 (Melanoma Antigen Recognised by T cells-1) MAGE-1, MAGE-3, 5T4, gp100, Carcinoembryonic antigen (CEA), prostate-specific antigen (PSA), MUCIN (MUC-1), tyrosinase. The TAA 5T4 (see WO 89/07947) has been extensively characterised. It is a 72kDa glycoprotein expressed widely in carcinomas, but having a highly restricted expression pattern in normal adult tissues. It appears to be strongly correlated to metastasis in colorectal and gastric cancer.

In the context of the present invention, TAAs are important since they may be used in ATC to deliver drugs to malignant cells but not most normal cells. For example, the highly specific antibody Mylotarg recognises a TAA called "CD33" which is abundant on AML cells but absent from normal blood stem cells. The antibody is linked to a novel and extremely potent chemotherapy agent known as "calicheamicin". The antibody selectively targets leukemic blast cells and delivers calicheamicin to them.

As a result, the leukemic cells are destroyed but the cells that are responsible for replenishing normal blood cells are spared.

DETERMINING THE UTILITY OF A NUCLEOTIDE SEQUENCE OF INTEREST AS A DISEASE TARGET

In one aspect, the present invention relates to a method for determining the utility of a nucleotide sequence of interest as a disease target – such as a TAA. A nucleotide sequence of interest is ligated into an expression vector according to the present invention – such as pIRES_STAR – to covalently join the nucleic acid sequence of interest to the expression vector. The expression vector is then transformed/transfected into a host cell (e.g. a mammalian cell - such as a CHO cell) and transformants comprising the expression vector are selected using methods well known in the art. The expression profile of the protein expressed from the nucleotide sequence of interest is then determined to establish if the nucleotide sequence of interest is a disease target – such as a TAA.

EXPRESSION PROFILE

As used herein, the term "expression profile" relates to identifying one or more of the following properties of the protein expressed from the nucleotide sequence of interest: (1) identifying the cellular location of the expressed protein; (2) determining if the expressed protein is internalised on binding an antibody; (3) determining if the expressed protein is secreted; (4) establishing the size of the expressed protein; and (5) determining the cytotoxic effects of an antibody binding to a cell expressing the tagged target antigen.

Preferably, the expression profile of the nucleotide sequence of interest is determined by: (1) identifying the cellular location of the expressed protein; (2) determining whether membrane-bound proteins are internalised on binding an antibody; (3) determining whether the protein is secreted; and (4) establishing the size of the gene product. Accordingly, this type of expression profile is also known as a LISS profile.

The expression profile of the protein expressed from the nucleotide sequence of interest can be determined as outlined below:

1. Identifying the cellular location of the expressed protein

In accordance with the present invention, the cellular location of the protein expressed from the nucleotide sequence of interest is identified in determining the expression profile. Preferably, the protein expressed from the nucleotide sequence of interest is associated with the cell-surface e.g. the plasma membrane.

In a preferred embodiment, flow cytometry of the binding of a labelled anti-amino-terminal tag antibody e.g. an anti-myc or anti-his antibody, is used to determine the expression profile of the nucleotide sequence of interest.

Flow cytometry is a powerful method for studying and purifying cells. It has found wide application, particularly in immunology and cell biology: however, the capabilities of the FACS can be applied in many other fields of biology. The acronym F.A.C.S. stands for Fluorescence Activated Cell Sorting, and is used interchangeably with "flow cytometry". The principle of FACS is that individual cells, held in a thin stream of fluid, are passed through one or more laser beams, causing light to be scattered and fluorescent dyes to emit light at various frequencies. Photomultiplier tubes (PMT) convert light to electrical signals, which are interpreted by software to generate data about the cells. Sub-populations of cells with defined characteristics can be identified and automatically sorted from the suspension at very high purity (~100%). For a general reference, see Flow Cytometry and Cell Sorting: A Laboratory Manual (1992) A. Radbruch (Ed.), Springer Laboratory, New York.

FACS machines collect fluorescence signals in one to several channels corresponding to different laser excitation and fluorescence emission wavelengths. Fluorescent labelling allows the investigation of many aspects of cell structure and function. The most widely used application is immunofluorescence: the staining of cells with

antibodies conjugated to fluorescent dyes e.g. fluorescein and phycoerythrin. This method is often used to label molecules on the cell surface, but antibodies can also be directed at targets within the cell. In direct immunofluorescence, an antibody to a particular molecule is directly conjugated to a fluorescent dye. Cells can then be stained in one step. In indirect immunofluorescence, the primary antibody is not labelled, but a second fluorescently conjugated antibody is added which is specific for the first antibody.

FACS may be used to measure gene expression in cells transfected with the expression vector of the present invention. This can be achieved directly, by labelling of the protein product, or indirectly by using a reporter gene in the construct.

Examples of reporter genes are -galactosidase and Green Fluorescent Protein (GFP). β-galactosidase activity can be detected by FACS using fluorogenic substrates - such as fluorescein digalactoside (FDG). FDG is introduced into cells by hypotonic shock, and is cleaved by the enzyme to generate a fluorescent product, which is trapped within the cell. One enzyme can therefore generate a large amount of fluorescent product. Cells expressing GFP constructs will fluoresce without the addition of a substrate. Mutants of GFP are available which have different excitation frequencies, but which emit fluorescence in the same channel. In a two-laser FACS machine, it is possible to distinguish cells, which are excited by the different lasers and therefore assay two transfections at the same time. FACS machines and reagents for use in FACS are widely available from sources world-wide including Becton-Dickinson, or Arizona Research Laboratories (http://www.arl.arizona.edu/facs/).

Preferably, cells are labeled with a primary antibody. For example, the primary antibody may bind the protein expressed from the nucleotide sequence of interest or the amino-terminal tag sequence and may be a mouse anti-c-Myc or a mouse anti-6-His mAb. The cells are then washed and labeled with a secondary antibody conjugated to, for example, rabbit anti-mouse-FITC.

Fluorescein (abbreviated by its commonly-used reactive isothiocyanate form, FITC) is currently the most commonly-used fluorescent dye for FACS analysis. FITC is a small organic molecule, and is typically conjugated to proteins via primary amines (i.e., lysines). Usually, between 3 and 6 FITC molecules are conjugated to each antibody; higher conjugations may result in solubility problems as well as internal quenching (and reduced brightness). Thus, an antibody will usually be conjugated in several parallel reactions to different amounts of FITC, and the resulting reagents will be compared for brightness (and background stickiness) to choose the optimal conjugation ratio. Fluorescein is typically excited by the 488 nm line of an argon laser, and emission is collected at 530 nm.

The cells are again washed, fixed in formalin and analysed. Dead /dying cells may be excluded from the analysis by selecting appropriate forward and side scatter populations.

Suitable imaging agents for use with FACS may be delivered to the cells by any suitable technique, including simple exposure thereto in cell culture, delivery of transiently expressing nucleic acids by viral or non- viral vector means, liposome-mediated transfer of nucleic acids or imaging agents, and the like.

Alternative means of cell sorting may also be employed. For example, nucleic acid probes complementary to the nucleotide sequence of interest may be used. Such probes can be used to identify cells expressing the nucleotide sequence of interest individually, such that they may subsequently be sorted either manually, or using FACS sorting. Nucleic acid probes complementary to the nucleotide sequence of interest mRNA may be prepared according to the teaching set forth above, using the general procedures as described by Sambrook et al (1989).

Fluorescence microscopy of the binding of a fluorescent labelled anti-amino-terminal tag antibody (for example, anti-myc or anti-his antibody) may also be used to allow determination of the cellular location of the gene product. Glass chamber slides may be coated with poly-D-lysine overnight, washed and cells seeded onto them. When the

appropriate cell density is achieved, cells are fixed in methanol/acetone, washed, incubated with primary antibody (e.g. an antibody to the protein expressed from the nucleotide sequence of interest, anti-Myc, or anti-6-His) washed, labeled with secondary antibody (e.g. rabbit-anti-mouse-FITC) washed and mounted prior to analysis by fluorescence microscopy.

2. Determining if the expressed protein is internalised on binding an antibody

Where it is demonstrated that the protein expressed from the nucleotide sequence of interest is associated with the cell-surface e.g. the plasma membrane, it is then necessary to determine whether addition of an antibody leads to internalisation of the protein expressed from the nucleotide sequence of interest.

Preferably, FACS analysis is used to determine whether the addition of an antibody – (e.g. an anti-myc antibody) leads to internalisation. Cells are labeled with primary antibody which may include a primary antibody that binds the protein expressed from the nucleotide sequence of interest or the amino-terminal tag sequence – such as mouse monoclonal anti-c-Myc or mouse monoclonal anti-6-His. The cells are washed and resuspended in tissue culture medium, and incubated or placed directly on ice, and after each time point, samples are washed and left overnight. Secondary antibody conjugated to, for example, rabbit anti-mouse-FITC is then added, cells are washed and flow cytometry is performed. TOPRO-3 may be added to exclude dead/dying cells from the analysis.

A person skilled in the art will understand that internalisation that may occur naturally due to ligand binding may not always be mimicked by binding of an antibody to the amino-terminal tag, but where this does occur it will demonstrate that the protein is capable of being internalised.

3. Determining if the expressed protein is secreted

Various methods known in the art may be used to detect if the protein expressed from the nucleotide sequence of interest is secreted and may include Western blotting and Enzyme Linked Immuno Sorbent Assay. Preferably, the protein expressed from the nucleotide sequence of interest is not secreted.

For Western analysis of supernatants from transfected cells, cells may be incubated in serum-free medium and supernatants centrifuged to remove cells, filtered and concentrated prior to denaturation (reducing or non-reducing conditions) and loading onto gels, subsequent electrophoresis, Western blotting and immunodetection of the protein expressed from the nucleotide sequence of interest or the amino-terminal tag, for example.

For Western analysis of cell lysates, cells may be scraped off into tissue culture medium, washed and resuspended in loading buffer. Following the shearing of cells through a needle and denaturation (reducing or non-reducing conditions), lysates are loaded onto gels and subjected to electrophoresis, Western blotting and immunodetection of the protein expressed from the nucleotide sequence of interest or the amino-terminal tag, for example.

Various methods for Western blotting are available including methods that are commercially available - such as ECL™ Western blotting from Amersham Pharmacia Biotech which is a light emitting non-radioactive method for detection of immobilised specific antigens, conjugated directly or indirectly with horseradish peroxidase-labelled antibodies.

If the protein expressed from the nucleotide sequence of interest is not secreted from the cell then the protein or the amino-terminal tag for example, will not be detected in the supernatants from transfected cells but will be detected in the cell lysates. If the protein expressed from the nucleotide sequence of interest is secreted from the cell then the protein or the amino-terminal tag for example, will be detected in the supernatants and the cell lysates from transfected cells.

4. Establishing the size of the expressed protein

Detection of secretion size may be carried out using various methods known to a person skilled in the art including Western blot analysis of supernatants and cell lysates from transfected cells. This analysis may provide further information about the expressed protein such as whether the expressed protein is glycosylated.

5. Determining the cytotoxic effects of an antibody binding to a cell expressing the tagged target antigen.

The cytotoxic effects of the anti-amino-terminal tag antibody binding to the target antigen may be assessed using an antibody to the tagged protein. It is apparent that many anti-cancer antibodies may work by non-immunlogical methods *e.g.* effecting receptor turn-over as for Herceptin. A cytoxic assay using an antibody to the tagged antigen may be used to test for this. By way of example, target cells expressing the tagged antigen are grown in 96 well tissue culture plates. Cells are seeded at a range of densities e.g. 100-5000 cells per well. Anti-tag antibody is added to the wells at a range of concentrations. Cells are incubated with label e.g. MMT, to monitor DNA replication. After specific incubation periods, cells are counted or harvested and DNA replication assessed and compared to controls.

In another aspect, the present invention provides a method for determining the utility of a nucleotide sequence of interest as a disease target, comprising the step of purifying the protein expressed from the nucleotide sequence of interest. Preferably, the protein expressed from the nucleotide sequence of interest is purified using the purification means contained within the amino-terminal tag, for example, 6-His.

PREPARING AN ANTIBODY

In a further aspect, the present invention provides a method for preparing an antibody or antibody fragment against a disease target polypeptide identified by the methods of

the present invention comprising the step of expressing the nucleotide sequence of interest encoding the disease target polypeptide, for example, a TAA polypeptide.

Antibodies according to the present invention may be produced in any suitable cell culture. This may be accomplished *via* bacterial, yeast, fungal, plant or mammalian cell culture. Preferably, mammalian cell culture is used. The selected cell culture system may secrete the antibody product.

The method comprises culturing a host, e.g. *E.coli*, yeast or a mammalian cell line, under conditions that allow for expression of the nucleotide sequence of interest encoding the disease target polypeptide. Preferably, the host cell is a transformed cell. Typically the host cell is prepared by transforming a cell with a vector comprising at least the nucleotide sequence of interest to be expressed in order to form the transformed cell. A typical vector for the transformation is an expression vector, for example, pIRES_STAR.

Multiplication of hybridoma cells or mammalian host cells *in vitro* is carried out in suitable culture media known to those skilled in the art, for example Dulbecco's Modified Eagle Medium (DMEM) or RPMI 1640 medium, optionally supplemented with a mammalian serum, e.g. foetal calf serum, or with trace elements and growth sustaining supplements, e.g. feeder cells such as normal mouse peritoneal exudate cells, spleen cells, bone marrow macrophages, 2-aminoethanol, insulin, transferrin, low density lipoprotein, oleic acid, or the like. Multiplication of bacterial or yeast host cells is carried out in suitable culture media known in the art, for example for bacteria in Luria Broth (LB), NZCYM, NZYM, NZM, Terrific Broth, SOB, SOC, 2 x YT, or M9 Minimal Medium, and for yeast in YPD, YEPD, Minimal Medium, or Complete Minimal Dropout Medium.

The cell culture supernatants are screened for the prescence of the desired antibodies using methods known in the art such as immunofluorescent staining, immunoblotting, enzyme immunoassay, e.g. a sandwich assay or a dot-blot assay, by a

radioimmunoassay, or by any other suitable technique known to those skilled in the art.

In vitro production provides relatively pure antibody preparations and allows scale-up to give large amounts of the desired antibodies. Techniques for bacterial cell, yeast or mammalian cell cultivation are known in the art and include homogeneous suspension culture, e.g. in an airlift reactor or in a continuous stirrer reactor, or immobilised or entrapped cell culture, e.g. in hollow fibres, microcapsules, on agarose microbeads or ceramic cartridges.

Large quantities of the desired antibodies can also be obtained by multiplying mammalian cells *in vivo*. For this purpose, hybridoma cells producing the desired antibodies are injected into histocompatible mammals to cause growth of antibody-producing tumours. Optionally, the animals are primed with a hydrocarbon, for example mineral oils such as pristane (tetramethyl-pentadecane), prior to the injection. After one to three weeks, the antibodies are isolated from the body fluids of those mammals.

The foregoing, and other, techniques are discussed in, for example, Kohler and Milstein, (1975) *Nature* 256:495-497; US 4,376,110; Harlow and Lane, Antibodies: a Laboratory Manual, (1988) Cold Spring Harbor Laboratory Press, incorporated herein by reference. Techniques for the preparation of recombinant antibody molecules are described in the above references and also in, for example, EP-A-0623679; EP-A-0368684 and EP-A-0436597, which are incorporated herein by reference.

For isolation of the antibodies, the immunoglobulins may be concentrated, for example by precipitation with ammonium sulphate, by dialysis against hygroscopic material such as polyethylene glycol, by filtration through selective membranes, or the like. Optionally, the antibodies may be purified by chromatography, for example gel filtration, ion-exchange chromatography, chromatography over DEAE-cellulose, or affinity chromatography. Affinity chromatography may be with the disease target – such as a TAA - or with a super-antigen such as Protein-A, Protein-G, Protein-L.

Preferably, the antibodies are monclonal antibodies. Thus, the present invention further concerns hybridoma cells secreting the monoclonal antibodies of the present invention. The preferred hybridoma cells of the invention are genetically stable and secrete monoclonal antibodies of the invention of the desired specificity.

A hybridoma cell line may be prepared that secretes monoclonal antibodies directed against the disease target polypeptides identified in the present invention. A suitable mammal – such as a mouse - is immunised with the purified disease target polypeptide, or with a carrier bearing a purified disease target polypeptide, or with cells bearing the disease target polypeptide. After an appropriate schedule of immunisation, known to those skilled in the art and taught in (Harlow and Lane, Antibodies: a Laboratory Manual, (1988) Cold Spring Harbor Laboratory Press), antibody-producing cells of the immunised mammal are fused with cells of a suitable myeloma cell line, the hybrid cells obtained in the fusion are cloned, and cell clones secreting the desired antibodies are selected. The resulting hybrid cells are screened for secretion of the desired antibodies, and those found to be expressing such antibodies are cloned.

NUCLEOTIDE SEQUENCE LIBRARY

Aspects of the present invention relate to nucleotide sequence libraries. As used herein, the term "nucleotide sequence library" refers to a physical library of nucleotide sequences of interest – such as disease targets and/or TAAs - ligated in to the expression vector of the present invention e.g. pIRES_STAR.

In another aspect, the present invention relates to a method for preparing nucleotide sequence libraries.

The physical library of nucleotide sequences of interest may be prepared by treating a nucleic acid sample – such as cDNA or genomic DNA - to fragment the nucleic acid therein. Preferably, the nucleic acid sample comprises cDNA. This is followed by

ligation of the nucleic acid fragments in to a vector – such as a cloning vector - using methods well know in the art, which covalently joins the nucleic acid fragments to the plasmid DNA. The vector is transformed into a host cell and transformants comprising the vector are selected using methods well known in the art. The signal sequence is then identified by sequencing the DNA inserted into the vector and using programs e.g. the SignalP program. The gene immediately downstream of the signal sequence is amplified and the nucleotide sequence of interest is subcloned into the expression vector of the present invention. The expression profile of the protein expressed from the nucleotide sequence of interest is then determined to establish if the nucleotide sequence of interest is a disease target – such as a TAA.

HOST CELLS

As used herein, the term "host cell" refers to any cell that comprises nucleotide sequences that are of use in the present invention.

Host cells may be transformed or transfected with a nucleotide sequence contained in a vector e.g. a cloning vector. Preferably said nucleotide sequence is carried in a vector for the replication and/or expression of the nucleotide sequence. The cells will be chosen to be compatible with the said vector and may for example be prokaryotic (for example bacterial), fungal, yeast or plant cells.

The gram-negative bacterium *E. coli* is widely used as a host for cloning nucleotide sequences. This organism is also widely used for heterologous nucleotide sequence expression. However, large amounts of heterologous protein tend to accumulate inside the cell. Subsequent purification of the desired protein from the bulk of *E. coli* intracellular proteins can sometimes be difficult.

In contrast to *E. coli*, bacteria from the genus Bacillus are very suitable as heterologous hosts because of their capability to secrete proteins into the culture medium. Other bacteria suitable as hosts are those from the genera Streptomyces and Pseudomonas.

Depending on the nature of the polynucleotide and/or the desirability for further processing of the expressed protein, eukaryotic hosts including yeasts or other fungi may be preferred. In general, yeast cells are preferred over fungal cells because yeast cells are easier to manipulate. However, some proteins are either poorly secreted from the yeast cell, or in some cases are not processed properly (*e.g.* hyperglycosylation in yeast). In these instances, a different fungal host organism should be selected.

Examples of expression hosts are fungi - such as Aspergillus species (such as those described in EP-A-0184438 and EP-A-0284603) and Trichoderma species; bacteria - such as Bacillus species (such as those described in EP-A-0134048 and EP-A-0253455), Streptomyces species and Pseudomonas species; and yeasts - such as Kluyveromyces species (such as those described in EP-A-0096430 and EP-A-0301670) and Saccharomyces species. By way of example, typical expression hosts may be selected from *Aspergillus niger*, *Aspergillus niger var. tubigenis*, *Aspergillus niger var. awamori*, *Aspergillus aculeatis*, *Aspergillus nidulans*, *Aspergillus orvzae*, *Trichoderma reesei*, *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Kluyveromyces lactis* and *Saccharomyces cerevisiae*.

The use of host cells - such as yeast, fungal and plant host cells - may provide for post-translational modifications (*e.g.* myristoylation, glycosylation, truncation, lapidation and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the present invention.

Aspects of the present invention also relate to host cells comprising the expression vector of the present invention. The expression vector may comprise a nucleotide sequence (e.g. a nucleotide sequence of interest) for replication and expression of the sequence. The cells will be chosen to be compatible with the vector and may for example be prokaryotic (for example bacterial), fungal, yeast or plant cells.

Preferably, the host cells are mammalian cells – such as CHO cells.

TRANSFECTION

Introduction of a vector into a host cell can be effected by various methods. For example, calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction or infection may be used. Such methods are described in many standard laboratory manuals - such as Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Host cells containing the expression vector can be selected by using, for example, G418 for cells transfected with an expression vector carrying a neomycin resistance selectable marker.

TRANSFORMATION

Teachings on the transformation of cells are well documented in the art, for example see Sambrook *et al* (Molecular Cloning: A Laboratory Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press) and Ausubel *et al.*, Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.

If a prokaryotic host is used then the nucleotide sequence may need to be suitably modified before transformation - such as by removal of introns.

A host cell may be transformed with a nucleotide sequence. Host cells transformed with the nucleotide sequence may be cultured under conditions suitable for the replication or expression of the nucleotide sequence.

CONSTRUCTS

Nucleotide sequences may be present in a construct.

The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes a nucleotide sequence directly or indirectly attached to a promoter. An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence including the Sh1-intron or the ADH intron, intermediate to the promoter and the nucleotide sequence. The same is true for the term "fused" which includes direct or indirect attachment. In some cases, the terms do not cover the natural combination of the nucleotide sequence coding for the protein ordinarily associated with the wild type nucleotide sequence promoter and when they are both in their natural environment.

The construct may even contain or express a marker, which allows for the selection of the nucleotide sequence construct in, for example, a bacterium, preferably of the genus Bacillus, such as *Bacillus subtilis*, or plants into which it has been transferred. Various markers exist which may be used, for example those encoding mannose-6-phosphate isomerase (especially for plants) or those markers that provide for antibiotic resistance - *e.g.* resistance to G418, hygromycin, bleomycin, kanamycin and gentamycin.

VECTORS

Nucleotide sequences may be present in a vector.

The term "vector" includes expression vectors and transformation vectors and shuttle vectors.

The term "transformation vector" means a construct capable of being transferred from one entity to another entity - which may be of the species or may be of a different species. If the construct is capable of being transferred from one species to another e.g. from an *E. coli* plasmid to a bacterium, such as of the genus Bacillus, then the transformation vector is sometimes called a "shuttle vector". It may even be a construct capable of being transferred from an *E. coli* plasmid to an Agrobacterium to a plant.

The vectors may be transformed into a suitable host cell as described below to provide for expression of a polypeptide.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter.

The vectors may contain one or more selectable marker nucleotide sequences. The most suitable selection systems for industrial micro-organisms are those formed by the group of selection markers which do not require a mutation in the host organism. Examples of fungal selection markers are the nucleotide sequences for acetamidase (*amd*S), ATP synthetase, subunit 9 (*oli*C), orotidine-5'-phosphate-decarboxylase (*pyr*A), phleomycin and benomyl resistance (benA). Examples of non-fungal selection markers are the bacterial G418 resistance nucleotide sequence (this may also be used in yeast, but not in filamentous fungi), the ampicillin resistance nucleotide sequence (*E. coli*), the neomycin resistance nucleotide sequence (*Bacillus*) and the *E. coli uid*A nucleotide sequence, coding for β-glucuronidase (GUS).

Vectors may be used *in vitro*, for example for the production of RNA or used to transfect or transform a host cell.

Thus, polynucleotides may be incorporated into a recombinant vector (typically a replicable vector), for example a cloning or expression vector. The vector may be used to replicate the nucleic acid in a compatible host cell.

VARIANTS/HOMOLOGUES/DERIVATIVES

The present invention encompasses the use of fragments, variants, homologues and derivatives.

The term "variant" is used to mean a naturally occurring polypeptide or nucleotide sequences which differs from a wild-type sequence.

The term "fragment" indicates that a polypeptide or nucleotide sequence comprises a fraction of a wild-type sequence. It may comprise one or more large contiguous sections of sequence or a plurality of small sections. The sequence may also comprise other elements of sequence, for example, it may be a fusion protein with another protein. Preferably the sequence comprises at least 50%, more preferably at least 65%, more preferably at least 80%, most preferably at least 90% of the wild-type sequence.

The term "homologue" means an entity having a certain homology with the subject amino acid sequences and the subject nucleotide sequences. Here, the term "homology" can be equated with "identity".

In the present context, a homologous sequence is taken to include an amino acid sequence, which may be at least 75, 85 or 90 % identical, preferably at least 95 or 98 % identical to the subject sequence. Typically, the homologues will comprise the same active sites etc. as the subject amino acid sequence. Although homology can also be considered in terms of similarity (*i.e.* amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity.

In the present context, a homologous sequence is taken to include a nucleotide sequence, which may be at least 75, 85 or 90% identical, preferably at least 95 or 98 % identical to the subject sequence. Typically, the homologues will comprise the same sequences that code for the active sites etc. as the subject sequence. Although homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity.

Homology comparisons may be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs can calculate % homology between two or more sequences.

% homology may be calculated over contiguous sequences, i.e. one sequence is aligned with the other sequence and each amino acid in one sequence is directly compared with the corresponding amino acid in the other sequence, one residue at a time. This is called an "ungapped" alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues.

Although this is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion will cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in % homology when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without penalising unduly the overall homology score. This is achieved by inserting "gaps" in the sequence alignment to try to maximise local homology.

However, these more complex methods assign "gap penalties" to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. "Affine gap costs" are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons. For example when using the GCG Wisconsin Bestfit package the default gap penalty for amino acid sequences is -12 for a gap and -4 for each extension.

Calculation of maximum % homology therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux *et al.*, 1984, *Nucleic Acids Research*

12:387). Examples of other software than can perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel *et al.*, 1999 ibid – Chapter 18), FASTA (Atschul *et al.*, 1990, J. Mol. Biol., 403-410) and the GENEWORKS suite of comparison tools. Both BLAST and FASTA are available for offline and online searching (see Ausubel *et al.*, 1999 ibid, pages 7-58 to 7-60). However, for some applications, it is preferred to use the GCG Bestfit program. A new tool, called BLAST 2 Sequences is also available for comparing protein and nucleotide sequence (see FEMS Microbiol Lett 1999 174(2): 247-50; FEMS Microbiol Lett 1999 177(1): 187-8).

Although the final % homology can be measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is the BLOSUM62 matrix - the default matrix for the BLAST suite of programs. GCG Wisconsin programs generally use either the public default values or a custom symbol comparison table if supplied (see user manual for further details). For some applications, it is preferred to use the public default values for the GCG package, or in the case of other software, the default matrix - such as BLOSUM62.

Once the software has produced an optimal alignment, it is possible to calculate % homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

The sequences may also have deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in a functionally equivalent substance. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the secondary binding activity of the substance is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino

acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine.

Conservative substitutions may be made, for example according to the Table below. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar - uncharged	C S T M
		N Q
	Polar - charged	D E
		K R
AROMATIC		H F W Y

The present invention also encompasses homologous substitution (substitution and replacement are both used herein to mean the interchange of an existing amino acid residue, with an alternative residue) may occur i.e. like-for-like substitution - such as basic for basic, acidic for acidic, polar for polar etc. Non-homologous substitution may also occur i.e. from one class of residue to another or alternatively involving the inclusion of unnatural amino acids - such as ornithine (hereinafter referred to as Z), diaminobutyric acid ornithine (hereinafter referred to as B), norleucine ornithine (hereinafter referred to as O), pyriylalanine, thienylalanine, naphthylalanine and phenylglycine.

Replacements may also be made by unnatural amino acids include; alpha* and alpha-disubstituted* amino acids, N-alkyl amino acids*, lactic acid*, halide derivatives of natural amino acids - such as trifluorotyrosine*, p-Cl-phenylalanine*, p-Br-phenylalanine*, p-I-phenylalanine*, L-allyl-glycine*, ß-alanine*, L-α-amino butyric acid*, L-γ-amino butyric acid*, L-α-amino isobutyric acid*, L-ε-amino caproic acid#,

7-amino heptanoic acid*, L-methionine sulfone#*, L-norleucine*, L-norvaline*, p-nitro-L-phenylalanine*, L-hydroxyproline#, L-thioproline*, methyl derivatives of phenylalanine (Phe) - such as 4-methyl-Phe*, pentamethyl-Phe*, L-Phe (4-amino)#, L-Tyr (methyl)*, L-Phe (4-isopropyl)*, L-Tic (1,2,3,4-tetrahydroisoquinoline-3-carboxyl acid)*, L-diaminopropionic acid # and L-Phe (4-benzyl)*. The notation * has been utilised for the purpose of the discussion above (relating to homologous or non-homologous substitution), to indicate the hydrophobic nature of the derivative whereas # has been utilised to indicate the hydrophilic nature of the derivative, #* indicates amphipathic characteristics.

Variant amino acid sequences may include suitable spacer groups that may be inserted between any two amino acid residues of the sequence including alkyl groups - such as methyl, ethyl or propyl groups - in addition to amino acid spacers - such as glycine or β-alanine residues. A further form of variation involves the presence of one or more amino acid residues in peptoid form will be well understood by those skilled in the art. For the avoidance of doubt, "the peptoid form" is used to refer to variant amino acid residues wherein the α-carbon substituent group is on the residue's nitrogen atom rather than the α-carbon. Processes for preparing peptides in the peptoid form are known in the art, for example Simon RJ *et al.*, PNAS (1992) 89(20), 9367-9371 and Horwell DC, Trends Biotechnol. (1995) 13(4), 132-134.

The nucleotide sequences for use in the present invention may include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones and/or the addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the nucleotide sequences of interest may be modified by any method available in the art. Such modifications may be carried out to enhance the *in vivo* activity or life span of nucleotide sequences useful in the present invention.

The present invention may also involve the use of nucleotide sequences that are complementary to the nucleotide sequences of interest or any derivative, fragment or derivative thereof. If the sequence is complementary to a fragment thereof then that sequence can be used as a probe to identify similar coding sequences in other organisms etc.

AGENT

In one aspect of the present invention, agents may be identified that bind to the disease targets – such as TAAs – identified by the methods of the present invention.

The agent may be an organic compound or other chemical. The agent may be a compound, which is obtainable from or produced by any suitable source, whether natural or artificial. The agent may be an amino acid molecule, a polypeptide, or a chemical derivative thereof, or a combination thereof. The agent may even be a polynucleotide molecule - which may be a sense or an anti-sense molecule, or an inhibitory RNA, or an antibody.

Preferably, the agent is an antibody. More preferably, the agent is a monoclonal antibody. Most preferably, the agent is a monoclonal humanised antibody.

Various strategies have been developed to produce monoclonal antibodies with human character which bypasses the need for an antibody-producing human cell line. For example, useful mouse monoclonal antibodies have been "humanised" by linking rodent variable regions and human constant regions (Winter, G. and Milstein, C. (1991) *Nature* 349, 293-299). This reduces the human anti-mouse immunogenicity of the antibody but residual immunogenicity is retained by virtue of the foreign V-region framework. Moreover, the antigen-binding specificity is essentially that of the murine donor. CDR-grafting and framework manipulation (EP 0239400) has improved and refined antibody manipulation to the point where it is possible to produce humanised murine antibodies which are acceptable for therapeutic use in humans. Humanised

antibodies may be obtained using other methods well known in the art (for example as described in US-A-239400).

The agents are preferably attached to an entity (e.g. an organic molecule) by a linker which may be a hydrolysable bifunctional linker.

The entity may be designed or obtained from a library of compounds, which may comprise peptides, as well as other compounds, such as small organic molecules.

By way of example, the entity may be a natural substance, a biological macromolecule, or an extract made from biological materials such as bacteria, fungi, or animal (particularly mammalian) cells or tissues, an organic or an inorganic molecule, a synthetic agent, a semi-synthetic agent, a structural or functional mimetic, a peptide, a peptidomimetic, a peptide cleaved from a whole protein, or a peptides synthesised synthetically (such as, by way of example, either using a peptide synthesizer or by recombinant techniques or combinations thereof, a recombinant agent, an antibody, a natural or a non-natural agent, a fusion protein or equivalent thereof and mutants, derivatives or combinations thereof.

Typically, the entity will be an organic compound. For some instances, the organic compounds will comprise two or more hydrocarbyl groups. Here, the term "hydrocarbyl group" means a group comprising at least C and H and may optionally comprise one or more other suitable substituents. Examples of such substituents may include halo-, alkoxy-, nitro-, an alkyl group, a cyclic group etc. In addition to the possibility of the substituents being a cyclic group, a combination of substituents may form a cyclic group. If the hydrocarbyl group comprises more than one C then those carbons need not necessarily be linked to each other. For example, at least two of the carbons may be linked *via* a suitable element or group. Thus, the hydrocarbyl group may contain hetero atoms. Suitable hetero atoms will be apparent to those skilled in the art and include, for instance, sulphur, nitrogen and oxygen. For some applications, preferably the entity comprises at least one cyclic group. The cyclic group may be a polycyclic group, such as a non-fused polycyclic group. For some applications, the

entity comprises at least the one of said cyclic groups linked to another hydrocarbyl group.

The entity may contain halo groups. Here, "halo" means fluoro, chloro, bromo or iodo.

The entity may contain one or more of alkyl, alkoxy, alkenyl, alkylene and alkenylene groups – which may be unbranched- or branched-chain.

The entity may comprise a chemotherapeutic drug (Dubowchik & Walker (1999) *Pharmacol. Ther.* 83, 67-123), a toxin – such as a cytotoxin - e.g. a potent plant and bacterial toxin (Brinkmann (2000) *In Vivo* 14, 21-27), radionuclides (Illidge & Johnson (2000) *Br. J. Haematol.* 108, 679-688) or combinations thereof.

It will be appreciated by those skilled in the art that the entity may be derived from a prodrug. Examples of prodrugs include certain protected group(s) which may not possess pharmacological activity as such, but may, in certain instances, be administered (such as orally or parenterally) and thereafter metabolised in the body to form an entity that is pharmacologically active.

Suitable pro-drugs may include, but are not limited to, Doxorubicin, Mitomycin, Phenol Mustard, Methotraxate, Antifolates, Chloramphenicol, Camptothecin, 5-Fluorouracil, Cyanide, Quinine, Dipyridamole and Paclitaxel. Agents (e.g. an antibody or a fragment thereof) that bind the disease target identified using the methods of the present invention may be chemically linked to an enzyme of interest. Alternatively, the conjugate can be a fusion protein produced by recombinant DNA techniques with the antibody variable region genes and the gene encoding the enzyme. Preferably, the prodrug should be non-toxic, resistant to the action of endogenous enzymes, and be converted into active drug only by the targeted enzyme. The selective activation of anticancer prodrugs by mAb-enzyme conjugates is reviewed in Senetr & Springer (2001) *Advanced Drug Delivery Reviews* 53, 247-264.

It will be further appreciated that certain moieties known as "pro-moieties", for example as described in "Design of Prodrugs" by H. Bundgaard, Elsevier, 1985 (the disclosured of which is hereby incorporated by reference), may be placed on appropriate functionalities of the agents. Such prodrugs are also included within the scope of the invention.

Preferably, the agent results in the death of cells comprising a disease target – such as a TAA - following internalisation of the agent.

The agent may be in the form of a pharmaceutically acceptable salt – such as an acid addition salt or a base salt – or a solvate thereof, including a hydrate thereof. For a review on suitable salts see Berge *et al*, J. Pharm. Sci., 1977, 66, 1-19.

The agent of the present invention may be capable of displaying other therapeutic properties.

The agent may be used in combination with one or more other pharmaceutically active agents.

If combinations of active agents are administered, then the combinations of active agents may be administered simultaneously, separately or sequentially.

CYTOTOXIN

The cytotoxin may include, but is not limited to, potent toxins, toxin subunits or even ribosome inactivating proteins that kill cells by inactivating protein synthesis or signal transduction.

The cytoxic component/s may be synthetic or naturally occurring molecules. Naturally occurring molecules may be derived from any natural source – such as from plants, viruses, fungi or bacteria.

Examples of cytoxic components are calicheamicin, kainic acid, cytotoxin vac A of *helicobacter pylori*, heat stable entertoxin (eg ST1 and ST2), verocytoxin, cytochalasin, anthracylcines such as danunorubicin, doxorubicicin and nogalamycin, cyclophosphamide, mustargen-nitrogen mustard, melphalan, chlorobucil, cis-platinum, carbo platinum, vincristine, VP-16, prednisome, tiamanolone, methotrextate, vinca alkaloids, the anthracyclines. flourocil, paclitaxel, ricin toxin, diptheria toxin, gelonin, saporin, pokeweek, neocarzinostatin, geldanamycin, maytansin and gerntuzumab ozogamicin.

A further example of a cytoxic component is *Pseudomonas* exotoxin. A truncated form of *Psuedomonas* Exotoxin A (ETA) may be modified by replacemant of domain I with scFv specific to CD7 to target leukemic T cells (Peipp et al. Cancer Res (2002) 62 (10) 2848-2855). scFv specific to CD22 may be used to target B cell malignancies (Brinkmann et al. Mol. Med. Today (1996) 2 (10), 439-46). This scFv may be further stabilised by dimerisation of the scFv fragment which results in enhanced efficacy against leukemias resistant to purine analogues (Kreitman et al. N. Engl. J. Med. (2001) 345 (4) 241-7). Replacement of *Pseudomonas* ETA domain I with a scFv specific to CD25 may also be used against CD25 positive haematological malignancies (Kreitman et al. J Clin. Oncol. (2000) 18 (8) 1622-36). The remaining domains of *Pseudomonas* ETA (domain II and III) may allow access of the toxin to intracellular sites where it is activated following internalisation with either CD7, CD22 or CD25.

Another toxic compound, maytanisoid DM1 has been used successfully in pre-clinical studies to treat colorectal or prostate cancers. Conjugation of DM1 with antibodies specific to the TAAs, a mucin – type glycoprotein (Can-Ag) and a prostate stem cell antigen (PSTA) resulted in anti-tumour activity against colorectal and prostate tumour xenografts respectively (Liu *et al.* Proc. Natl. Acad. Sci. USA (1996) 93 (16) 8618-8623); Ross et al. Cancer Res. (2002) 62 (9) 2546-53).

The cytotoxin may even comprise an enzyme that activates a pro-drug, for example, carboxypeptidase G2, which may be used in Antibody directed Enzyme Prodrug Therapy (ADEPT). The enzyme may be fused to either of the antibodies specific to

5T4 (H8 or scFv). The enzyme-antibody conjugate may be administered first and any non-bound conjugate may then be allowed to clear before administration of the prodrug, 4[N,N-bis (2-iodoethyl) amino] phenoxycarbonyl L-glutamic acid (ZD2767P). This prodrug may be activated at either intracellular or extracellular sites to mediate its cytotoxic effects. This has been successfully demonstrated with the F(ab')2 fragment of CEA antibody A5B7 (Pedley et al. Cancer Res. (1999) 59 (16) 3998-4003).

The cytotoxin may even comprise a radiopharmaceutcial – such as alpha emitters ^{213}bismuth and ^{211}astatine, or ^{131}iodine (a mixed beta and gamma emitter), or ^{90}yttrium (a pure beta emiiter).

The cytotoxin may even include combinations of toxins - such as a combination of the *Pseudomonas* exotoxin and diptheria toxin (*Clin. Cancer Res.* (2002) 6, 942-944; *Immunol. Res.* (2002) 25, 177-191). Advantageously, such combinations may reduce or overcome problems associated with toxicity and/or immunogenicity.

The cytotoxin may even be a virus – such as a parvovirus.

One skilled in the art will appreciate that this list is not intended to be exhaustive and will be aware of other suitable cytotoxic components.

STEREO AND GEOMETRIC ISOMERS

The entity may exist as stereoisomers and/or geometric isomers – e.g. the entity may possess one or more asymmetric and/or geometric centres and so may exist in two or more stereoisomeric and/or geometric forms. The present invention contemplates the use of all the individual stereoisomers and geometric isomers of those entities, and mixtures thereof.

PHARMACEUTICAL SALT

The agents of the present invention may be administered in the form of a pharmaceutically acceptable salt.

Pharmaceutically-acceptable salts are well known to those skilled in the art, and for example include those mentioned by Berge *et al*, in J.Pharm.Sci., 66, 1-19 (1977). Suitable acid addition salts are formed from acids which form non-toxic salts and include the hydrochloride, hydrobromide, hydroiodide, nitrate, sulphate, bisulphate, phosphate, hydrogenphosphate, acetate, trifluoroacetate, gluconate, lactate, salicylate, citrate, tartrate, ascorbate, succinate, maleate, fumarate, gluconate, formate, benzoate, methanesulphonate, ethanesulphonate, benzenesulphonate and p-toluenesulphonate salts.

When one or more acidic moieties are present, suitable pharmaceutically acceptable base addition salts can be formed from bases which form non-toxic salts and include the aluminium, calcium, lithium, magnesium, potassium, sodium, zinc, and pharmaceutically-active amines such as diethanolamine, salts.

A pharmaceutically acceptable salt of an agent may be readily prepared by mixing together solutions of the agent and the desired acid or base, as appropriate. The salt may precipitate from solution and be collected by filtration or may be recovered by evaporation of the solvent.

The agent of the present invention may exist in polymorphic form.

The agent of the present invention may contain one or more asymmetric carbon atoms and therefore exists in two or more stereoisomeric forms. Where an agent contains an alkenyl or alkenylene group, cis (E) and trans (Z) isomerism may also occur. The present invention includes the individual stereoisomers of the agent and, where appropriate, the individual tautomeric forms thereof, together with mixtures thereof.

Separation of diastereoisomers or cis and trans isomers may be achieved by conventional techniques, e.g. by fractional crystallisation, chromatography or H.P.L.C.

of a stereoisomeric mixture of the agent or a suitable salt or derivative thereof. An individual enantiomer of the agent may also be prepared from a corresponding optically pure intermediate or by resolution, such as by H.P.L.C. of the corresponding racemate using a suitable chiral support or by fractional crystallisation of the diastereoisomeric salts formed by reaction of the corresponding racemate with a suitable optically active acid or base, as appropriate.

The agent may also include all suitable isotopic variations of the agent or a pharmaceutically acceptable salt thereof. An isotopic variation of an agent or a pharmaceutically acceptable salt thereof is defined as one in which at least one atom is replaced by an atom having the same atomic number but an atomic mass different from the atomic mass usually found in nature. Examples of isotopes that can be incorporated into the agent and pharmaceutically acceptable salts thereof include isotopes of hydrogen, carbon, nitrogen, oxygen, phosphorus, sulphur, fluorine and chlorine such as ^{2}H, ^{3}H, ^{13}C, ^{14}C, ^{15}N, ^{17}O, ^{18}O, ^{31}P, ^{32}P, ^{35}S, ^{18}F and ^{36}Cl, respectively. Certain isotopic variations of the agent and pharmaceutically acceptable salts thereof, for example, those in which a radioactive isotope such as ^{3}H or ^{14}C is incorporated, are useful in drug and/or substrate tissue distribution studies. Tritiated, i.e., ^{3}H, and carbon-14, i.e., ^{14}C, isotopes are particularly preferred for their ease of preparation and detectability. Further, substitution with isotopes such as deuterium, i.e., ^{2}H, may afford certain therapeutic advantages resulting from greater metabolic stability, for example, increased *in vivo* half-life or reduced dosage requirements and hence may be preferred in some circumstances. Isotopic variations of the agent and pharmaceutically acceptable salts thereof of this invention can generally be prepared by conventional procedures using appropriate isotopic variations of suitable reagents.

PHARMACEUTICALLY ACTIVE SALT

The agent may be administered as a pharmaceutically acceptable salt. Typically, a pharmaceutically acceptable salt may be readily prepared by using a desired acid or base, as appropriate. The salt may precipitate from solution and be collected by filtration or may be recovered by evaporation of the solvent.

CHEMICAL SYNTHESIS METHODS

The agent may be prepared by chemical synthesis techniques.

It will be apparent to those skilled in the art that sensitive functional groups may need to be protected and deprotected during synthesis of a compound of the invention. This may be achieved by conventional techniques, for example as described in "Protective Groups in Organic Synthesis" by T W Greene and P G M Wuts, John Wiley and Sons Inc. (1991), and by P.J.Kocienski, in "Protecting Groups", Georg Thieme Verlag (1994).

It is possible during some of the reactions that any stereocentres present could, under certain conditions, be racemised, for example if a base is used in a reaction with a substrate having an having an optical centre comprising a base-sensitive group. This is possible during e.g. a guanylation step. It should be possible to circumvent potential problems such as this by choice of reaction sequence, conditions, reagents, protection/deprotection regimes, etc. as is well-known in the art.

The compounds and salts may be separated and purified by conventional methods.

Separation of diastereomers may be achieved by conventional techniques, e.g. by fractional crystallisation, chromatography or H.P.L.C. of a stereoisomeric mixture of a compound of formula (I) or a suitable salt or derivative thereof. An individual enantiomer of a compound of formula (I) may also be prepared from a corresponding optically pure intermediate or by resolution, such as by H.P.L.C. of the corresponding racemate using a suitable chiral support or by fractional crystallisation of the diastereomeric salts formed by reaction of the corresponding racemate with a suitably optically active acid or base.

The agent or variants, homologues, derivatives, fragments or mimetics thereof may be produced using chemical methods to synthesise the agent in whole or in part. For example, if the agent comprises a peptide, then the peptide can be synthesised by solid

phase techniques, cleaved from the resin, and purified by preparative high performance liquid chromatography (e.g., Creighton (1983) Proteins Structures And Molecular Principles, WH Freeman and Co, New York NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, *supra*).

Synthesis of peptide inhibitor agents (or variants, homologues, derivatives, fragments or mimetics thereof) can be performed using various solid-phase techniques (Roberge JY *et al* (1995) Science 269: 202-204) and automated synthesis may be achieved, for example, using the ABI 43 1 A Peptide Synthesizer (Perkin Elmer) in accordance with the instructions provided by the manufacturer. Additionally, the amino acid sequences comprising the agent, may be altered during direct synthesis and/or combined using chemical methods with a sequence from other subunits, or any part thereof, to produce a variant agent.

CHEMICAL DERIVATIVE

The term "derivative" or "derivatised" as used herein includes chemical modification of an agent. Illustrative of such chemical modifications would be replacement of hydrogen by a halo group, an alkyl group, an acyl group or an amino group.

CHEMICAL MODIFICATION

The agent may be a chemically modified agent.

The chemical modification of an agent may either enhance or reduce hydrogen bonding interaction, charge interaction, hydrophobic interaction, Van Der Waals interaction or dipole interaction between the agent and the disease target – such as a TAA.

In one aspect, the agent may act as a model (for example, a template) for the development of other compounds.

PHARMACEUTICAL COMPOSITIONS

Pharmaceutical compositions of the present invention may comprise a therapeutically effective amount of the agent.

The pharmaceutical compositions may be for human or animal usage in human and veterinary medicine and will typically comprise any one or more of a pharmaceutically acceptable diluent, carrier, or excipient. Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (A. R. Gennaro edit. 1985). The choice of pharmaceutical carrier, excipient or diluent can be selected with regard to the intended route of administration and standard pharmaceutical practice. The pharmaceutical compositions may comprise as - or in addition to - the carrier, excipient or diluent any suitable binder(s), lubricant(s), suspending agent(s), coating agent(s), solubilising agent(s).

Preservatives, stabilizers, dyes and even flavoring agents may be provided in the pharmaceutical composition. Examples of preservatives include sodium benzoate, sorbic acid and esters of p-hydroxybenzoic acid. Antioxidants and suspending agents may be also used.

There may be different composition/formulation requirements dependent on the different delivery systems. By way of example, the pharmaceutical composition of the present invention may be formulated to be administered using a mini-pump or by a mucosal route, for example, as a nasal spray or aerosol for inhalation or ingestable solution, or parenterally in which the composition is formulated by an injectable form, for delivery, by, for example, an intravenous, intramuscular or subcutaneous route. Alternatively, the formulation may be designed to be administered by a number of routes.

If the agent is to be administered mucosally through the gastrointestinal mucosa, it should be able to remain stable during transit though the gastrointestinal tract; for

example, it should be resistant to proteolytic degradation, stable at acid pH and resistant to the detergent effects of bile.

Where appropriate, the pharmaceutical compositions may be administered by inhalation, in the form of a suppository or pessary, topically in the form of a lotion, solution, cream, ointment or dusting powder, by use of a skin patch, orally in the form of tablets containing excipients such as starch or lactose, or in capsules or ovules either alone or in admixture with excipients, or in the form of elixirs, solutions or suspensions containing flavouring or colouring agents, or the pharmaceutical compositions can be injected parenterally, for example intravenously, intramuscularly or subcutaneously. For parenteral administration, the compositions may be best used in the form of a sterile aqueous solution which may contain other substances, for example enough salts or monosaccharides to make the solution isotonic with blood. For buccal or sublingual administration the compositions may be administered in the form of tablets or lozenges which can be formulated in a conventional manner.

The agents may be used in combination with a cyclodextrin. Cyclodextrins are known to form inclusion and non-inclusion complexes with drug molecules. Formation of a drug-cyclodextrin complex may modify the solubility, dissolution rate, bioavailability and/or stability property of a drug molecule. Drug-cyclodextrin complexes are generally useful for most dosage forms and administration routes. As an alternative to direct complexation with the drug the cyclodextrin may be used as an auxiliary additive, e.g. as a carrier, diluent or solubiliser. Alpha-, beta- and gamma-cyclodextrins are most commonly used and suitable examples are described in WO-A-91/11172, WO-A-94/02518 and WO-A-98/55148.

If the agent is a protein, then said protein may be prepared *in situ* in the subject being treated. In this respect, nucleotide sequences encoding said protein may be delivered by use of non-viral techniques (e.g. by use of liposomes) and/or viral techniques (e.g. by use of retroviral vectors) such that the said protein is expressed from said nucleotide sequence.

ADMINISTRATION

The term "administered" includes delivery by viral or non-viral techniques. Viral delivery mechanisms include but are not limited to adenoviral vectors, adeno-associated viral (AAV) vectos, herpes viral vectors, retroviral vectors, lentiviral vectors, and baculoviral vectors. Non-viral delivery mechanisms include lipid mediated transfection, liposomes, immunoliposomes, lipofectin, cationic facial amphiphiles (CFAs) and combinations thereof.

The components may be administered alone but will generally be administered as a pharmaceutical composition – e.g. when the components are is in admixture with a suitable pharmaceutical excipient, diluent or carrier selected with regard to the intended route of administration and standard pharmaceutical practice.

For example, the components can be administered in the form of tablets, capsules, ovules, elixirs, solutions or suspensions, which may contain flavouring or colouring agents, for immediate-, delayed-, modified-, sustained-, pulsed- or controlled-release applications.

If the pharmaceutical is a tablet, then the tablet may contain excipients such as microcrystalline cellulose, lactose, sodium citrate, calcium carbonate, dibasic calcium phosphate and glycine, disintegrants such as starch (preferably corn, potato or tapioca starch), sodium starch glycollate, croscarmellose sodium and certain complex silicates, and granulation binders such as polyvinylpyrrolidone, hydroxypropylmethylcellulose (HPMC), hydroxypropylcellulose (HPC), sucrose, gelatin and acacia. Additionally, lubricating agents such as magnesium stearate, stearic acid, glyceryl behenate and talc may be included.

Solid compositions of a similar type may also be employed as fillers in gelatin capsules. Preferred excipients in this regard include lactose, starch, a cellulose, milk sugar or high molecular weight polyethylene glycols. For aqueous suspensions and/or elixirs, the agent may be combined with various sweetening or flavouring agents,

colouring matter or dyes, with emulsifying and/or suspending agents and with diluents such as water, ethanol, propylene glycol and glycerin, and combinations thereof.

The routes for administration (delivery) may include, but are not limited to, one or more of oral (e.g. as a tablet, capsule, or as an ingestable solution), topical, mucosal (e.g. as a nasal spray or aerosol for inhalation), nasal, parenteral (e.g. by an injectable form), gastrointestinal, intraspinal, intraperitoneal, intramuscular, intravenous, intrauterine, intraocular, intradermal, intracranial, intratracheal, intravaginal, intracerebroventricular, intracerebral, subcutaneous, ophthalmic (including intravitreal or intracameral), transdermal, rectal, buccal, vaginal, epidural, sublingual.

DOSE LEVELS

Typically, a physician will determine the actual dosage which will be most suitable for an individual subject. The specific dose level and frequency of dosage for any particular patient may be varied and will depend upon a variety of factors including the activity of the specific compound employed, the metabolic stability and length of action of that compound, the age, body weight, general health, sex, diet, mode and time of administration, rate of excretion, drug combination, the severity of the particular condition, and the individual undergoing therapy.

FORMULATION

The component(s) may be formulated into a pharmaceutical composition, such as by mixing with one or more of a suitable carrier, diluent or excipient, by using techniques that are known in the art.

SCREENING METHODS

A plurality of candidate compounds may be screened using the methods described below. In particular, these methods may be suited for screening libraries of

compounds including libraries of compounds that bind the disease targets – such as TAAs - identified by the methods of the present invention.

Where the candidate compounds are proteins e.g. antibodies or peptides, libraries of candidate compounds may be screened using phage display techniques. Phage display is a protocol of molecular screening, which utilises recombinant bacteriophage. The technology involves transforming bacteriophage with a gene that encodes the library of candidate compounds, such that each phage or phagemid expresses a particular candidate compound. The transformed bacteriophage (which preferably is tethered to a solid support) expresses the appropriate candidate compound and displays it on their phage coat. Specific candidate compounds which are capable of interacting with the disease targets identified by the methods of the present invention are enriched by selection strategies based on affinity interaction. The successful candidate agents are then characterised. Phage display has advantages over standard affinity ligand screening technologies. The phage surface displays the candidate agent in a three dimensional configuration, more closely resembling its naturally occurring conformation. This allows for more specific and higher affinity binding for screening purposes.

Another method of screening a library of compounds utilises eukaryotic or prokaryotic host cells, which are stably transformed with recombinant DNA molecules expressing the library of compounds. Such cells, either in viable or fixed form, can be used for standard binding-partner assays. See also Parce *et al.* (1989) Science 246:243-247; and Owicki *et al.* (1990) Proc. Nat'l Acad. Sci. USA 87;4007-4011, which describe sensitive methods to detect cellular responses. Competitive assays are particularly useful, where the cells expressing the library of compounds are incubated with a labelled antibody, such as ^{125}I-antibody, and a test sample such as a candidate compound whose binding affinity to the binding composition is being measured. The bound and free labelled binding partners are then separated to assess the degree of binding. The amount of test sample bound is inversely proportional to the amount of labelled antibody bound.

Any one of numerous techniques can be used to separate bound from free binding partners to assess the degree of binding. This separation step could typically involve a procedure such as adhesion to filters followed by washing, adhesion to plastic following by washing, or centrifugation of the cell membranes.

Still another approach is to use solubilized, unpurified or solubilized purified protein either extracted from expressing mammalian cells or from transformed eukaryotic or prokaryotic host cells. This allows for a "molecular" binding assay with the advantages of increased specificity, the ability to automate, and high drug test throughput.

Another technique for candidate compound screening involves an approach which provides high throughput screening for new compounds having suitable binding affinity and is described in detail in WO 84/03564. First, large numbers of different small peptide test compounds are synthesised on a solid substrate, *e.g.*, plastic pins or some other appropriate surface; see Fodor *et al.* (1991). Then all the pins are reacted with solubilised protein and washed. The next step involves detecting bound protein. Detection may be accomplished using a mAb to the protein of interest. Compounds which interact specifically with the protein may thus be identified.

Rational design of candidate compounds likely to be able to interact with the protein may be based upon structural studies of the molecular shapes of the protein and/or its *in vivo* binding partners. One means for determining which sites interact with specific other proteins is a physical structure determination, *e.g.*, X-ray crystallography or two-dimensional NMR techniques. These will provide guidance as to which amino acid residues form molecular contact regions. For a detailed description of protein structural determination, see, *e.g.*, Blundell and Johnson (1976) Protein Crystallography, Academic Press, New York.

GENERAL RECOMBINANT DNA METHODOLOGY TECHNIQUES

The present invention employs, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA and immunology, which are within the capabilities of a person of ordinary skill in the art. Such techniques are explained in the literature. See, for example, J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Books 1-3, Cold Spring Harbor Laboratory Press; Ausubel, F. M. et al. (1995 and periodic supplements; *Current Protocols in Molecular Biology*, ch. 9, 13, and 16, John Wiley & Sons, New York, N.Y.); B. Roe, J. Crabtree, and A. Kahn, 1996, *DNA Isolation and Sequencing: Essential Techniques*, John Wiley & Sons; M. J. Gait (Editor), 1984, *Oligonucleotide Synthesis: A Practical Approach*, Irl Press; and, D. M. J. Lilley and J. E. Dahlberg, 1992, *Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of DNA* Methods in Enzymology, Academic Press. Each of these general texts is herein incorporated by reference.

The invention will now be further described by way of Examples, which are meant to serve to assist one of ordinary skill in the art in carrying out the invention and are not intended in any way to limit the scope of the invention.

EXAMPLES

Example 1

Construction of pIRES_STAR

Using methods known in the art, the vector pIRES_STAR is constructed by inserting into pIRESneo (Genbank Accession Number U89673), a double stranded oligonucleotide comprising the sequence:

Seq ID No. 6

5'cgatccaccatgggatggagctgtatcatcctcttcttggtagcaacagctacaggtgtccactcccatcatcaccatcacc atgaacaaaaactcatctcagaagaggatctggcgtttaaacgcgatatcggcgcgccgc3'

This sequence encodes the murine Igκ signal sequence followed by 6 histidines, a myc epitope and four different 8 base recognition restriction enzyme sites, as illustrated in Figure 1. The oligonucleotide shown as Seq ID No. 4 is made from two complementary single stranded oligos that are annealed to give overhanging ends that are compatible with the *Cla*I and *Not*I sites of pIRESneo.

Thus, it is demonstrated that an expression vector in accordance with the present invention is constructed.

Example 2

Cloning Strategy

Novel cDNAs are obtained in standard cloning vectors. From the sequence, the leader sequence is identified using the Signal P program (http://www.cbs.dtu.dk/services/SignalP/). The method is described in Nielsen *et al.* (1997) *Protein Engineering* 10, 1-6. Briefly, a name for the sequence is optionally entered and the sequence is entered in the sequence window. The sequence must be written using the one letter amino acid code. It is recommend that the N-terminal part only of the protein is entered (not more than 50-70 amino acids) since longer sequences will increase the risk of false positives and make the graphical output difficult to read. One or more groups of organisms are chosen for the prediction by clicking the check-box next to the group(s): gram-: Use networks trained on sequences from gram-negative prokaryotes; gram+: Use networks trained on sequences from gram-positive prokaryotes; euk: Use networks trained on sequences from eukaryotes. A graphical output (in Postscript format) of the prediction is made available, if the "Include graphics" button is checked. The sequence is submitted and a WWW page returns the results when the prediction is ready.

A specific PCR primer is then designed to amplify the gene immediately downstream from this sequence that can be used in conjunction with a generic primer located

downstream of the cloning site in the vector. This gives a product representing the coding sequence of the protein product, incorporating its own termination codon, but minus the native leader sequence.

Normally, PCR amplified novel cDNAs are directionally cloned using the blunt *Pme*I site at the 5' end and the overhanging *Asc*I at the 3' end of the oligonucleotide cloned into pIRESneo. Thus, the 5' primer will not require an additional restriction site as *Pfu* polymerase leaves a blunt ended product, whilst the 3' primer will incorporate an additional *Asc*I site. Other restriction sites are included in the pIRES_STAR vector to allow alternative strategies if required.

Thus it is demonstrated that a nucleotide sequence of interest is obtained and cloned into an expression vector of the present invention.

Example 3

STAR-evaluation of 5T4

5T4 (see WO 89/07947) is a TAA that has been extensively characterised. It is a 72 kDa trophoblast glycoprotein frequently overexpressed in epithelial malignancies, but having a highly restricted expression pattern in normal adult tissues, providing a potential target for cancer therapeutics. The full nucleic acid sequence of human 5T4 is known (Myers *et al.*, 1994 *J Biol Chem* **169**: 9319-24).

(i) Materials & Methods

Preparation of constructs

Constructs for a membrane-bound protein and a secreted protein are made in which the tag is placed at the amino-terminus (pIRES_STARh5T4). Primers are designed to amplify a full length human 5T4 gene minus the signal sequence and human 5T4 lacking the transmembrane domain and cytoplasmic tail minus the signal sequence

(Figure 2). These constructs are then cloned into the pIRES_STAR expression vector (Figure 1). The expression profile of these proteins is compared with that of a previously made 5T4 protein in which a Myc-His tag is placed at the carboxy-terminus (pIRES_h5T4EKMH). The latter product has previously been shown to be expressed on the cell surface and is recognised by a 5T4 mAb (H8) that binds to a conformational epitope. A carboxy-terminally tagged Tm- h5T4 has been previously demonstrated to be secreted from mammalian cells.

Using the h5T4 gene with an optimised Kozak sequence (Kh5T4) obtained from Cytomix as target DNA, PCR is carried out using *Pfu* polmerase (Stratagene). The expected products are obtained. These are digested with *Asc*I, gel purified and then ligated in to pIRES_STAR, previously digested with *Pme*I and *Asc*I.

The constructs are transiently transfected into CHO cells and expression of h5T4 detected by immuno-staining of fixed cells with an anti-myc antibody followed by rabbit anti-mIg-HRP and detected with DAB. The anti h5T4 mAb, H8 is also used to detect the expressed protein (Figure 3) which demonstrates that the full length protein is expressed by the STAR expression vector as detected with the anti-Myc antibody.

Production of stable CHO STARh5T4 cell lines

Following transfection of CHO cells with pIRES_STARh5T4 or pIRES_STARh5T4Tm-, transfectants are selected using 1 mg/ml G418, and a number of ring clones are selected and propagated, always selecting cells immunostaining positive for anti-Myc and H8.

Antibodies

Mouse monoclonal anti-h5T4 H8 (purified antibody is used in all experiments except immunostaining, where hybridoma supernatant is used), mouse monoclonal anti-c-Myc (9E10 epitope), mouse monoclonal anti-PentaHis, rabbit anti-mouse-HRP and rabbit anti-mouse-FITC is used.

Immunostaining

Cells are fixed and permeabilised in methanol/acetone, washed and incubated with primary antibody, washed, incubated with secondary antibody conjugated to HRP, washed and DAB is added to visualize cells positive for h5T4 and the MycHis tag

Western analysis of supernatants from CHO_STARh5T4TM- cells

Cells are incubated in serum-free medium for 24h and supernatants are centrifuged to remove cells, filtered and concentrated prior to denaturation (reducing or non-reducing conditions) and loading onto gels, subsequent electrophoresis, Western blotting, immunodetection and ECL.

Western analysis of CHO_STARh5T4 cell lysates

Cells are scraped off into tissue culture medium, washed and resuspended in loading buffer. Following shearing of cells through a needle and denaturation (reducing or non-reducing conditions), lysates are loaded onto gels and subjected to electrophoresis, Western blotting and immunodetection/ECL as above.

FACS analysis

Cells are labeled with primary antibody, washed, labeled with secondary antibody conjugated to FITC, washed, fixed in formalin and analysed by flow cytometry. Dead /dying cells are excluded from the analysis by selecting appropriate forward and side scatter populations.

Internalisation assay by FACS

Cells are labeled with primary antibody, washed and resuspended in tissue culture medium, and incubated at 37 °C for 2 h or placed directly on ice, and after each time point, samples are washed at 4 °C and left at this temperature overnight and throughout

the rest of the procedure. Secondary antibody conjugated to FITC is then added, cells washed and flow cytometry is performed, with the addition of TOPRO-3 to exclude dead/dying cells from the analysis.

Immunocytochemistry

Glass chamber slides are coated with poly-D-lysine overnight at 37 °C, washed and cells are seeded onto them. When the appropriate cell density is achieved, cells are fixed in methanol/acetone, washed, incubated with primary antibody, washed, labeled with secondary antibody, washed and mounted prior to analysis by fluorescence microscopy.

(b) Results & Conclusions

Location

Figure 3 shows that the pattern of staining of permeabilised CHO_STARh5T4 is the same as for CHO_h5T4EKMH, *i.e.* the STARh5T4 protein is expressed in the membrane as seen by the stain being even over the cell surface. This is corroborated by the immunocytochemical data illustrated in Figure 4. FACS data illustrates in Figure 5A c) and d) that when cells are not permeabilised, H8, anti-Myc and anti-PentaHis can detect STARh5T4 on the surface of CHO_STARh5T4, as seen by the histogram peak shifts to the right compared to the negative control, thus signifying that the N-terminal MycHis tag and the h5T4 are extracellular. This is not the case with CHO_h5T4EKMH cells. Only H8 stains these as the C-terminal MycHis tag is inaccessible (intracellular)- see Figure 5A a) and b). The magnitude of peak shifts is summarised in Table 1.

Table 1: The geometric mean values calculated from histograms in Figure 5A for CHO_h5T4EKMH and CHO_STARh5T4 cells when different primary antibodies are used in FACS analysis.

Cell line	Antibody			
	Control (FACS buffer)	H8 (1:200)	Anti-Myc (1:500)	Anti-pentaHis (1:500)
CHO_h5T4EKMH	4.28	1407.66	5.70	8.14
CHO_STARh5T4 (clone 4)	2.75	1471.70	54.74	212.79

Internalisation

Figure 5Ba) shows an example of a cell line that can internalise h5T4 in response to H8, as seen by the shift of the histogram peak to the left (reduction in expression) in 2 h at 37 °C, representing antigen loss at the cell surface. Similarly in b), H8 can internalise h5T4 on the surface of CHO_h5T4EKMH, however, as seen in Figure 5A, there is no His expression extracellularly and as a result, anti-His cannot internalise h5T4EKMH over 2 h at 37 °C (Figure 5Bc). Most importantly, both H8 and anti-pentaHis can internalise STARh5T4 on CHO_STARh5T4 cells over the 2h period (Figure 5Bd) and e)). These data are summarised in Table 2.

Table 2: The geometric mean values calculated from histograms in Figure 5B for CHO_STARh5T4 incubated at 37 °C for t=0h and t=2h when H8 or anti-pentaHis antibodies are used, a) CHO_h5T4EKMH cells b) CHO_STARh5T4 (clone 4) cells.

a)

Geometric mean fluorescence (U)			
Antibody	Control (FACS buffer)	Time at 37^0C (h)	
		0	2
H8 (1:1000)	2.24	753.79	478.36
Anti-PentaHis (1:100)	2.24	3.51	3.25

b)

Geometric mean fluorescence (U)			
Antibody	Control (FACS buffer)	Time at 37^{0}C (h)	
		0	2
H8 (1:1000)	2.25	944.03	500.35
Anti-PentaHis (1:100)	2.25	61.11	24.06

Secretion

Western analysis as illustrated in Figure 6a) and b) shows that concentrated supernatant from CHO_STARh5T4 cells does not contain any detectable h5T4 or Myc containing protein, showing that these cells do not secrete STARh5T4, in contrast to CHO_STARh5T4TM -(clone 9) cells which secretes STARh5T4TM- as detected by anti-Myc and H8.

Size

Western analysis illustrated in Figure 6c) and d) shows that the STARh5T4 protein is similar in size to h5T4EKMH, *ie*. approximately 70kDa.

All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention

which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

CLAIMS

1. An expression vector comprising an amino-terminal tag sequence and a signal sequence operably linked to a nucleotide sequence of interest, wherein the amino-terminal tag sequence is inserted between the signal sequence and the nucleotide sequence of interest.

2. An expression vector according to claim 1 wherein the nucleotide sequence of interest is a disease target.

3. An expression vector according to claim 1 wherein the nucleotide sequence of interest is a tumour associated antigen.

4. An expression vector according to any one of claims 1 to 3 wherein the amino-terminal tag sequence comprises a 6-His sequence.

5. An expression vector according to any one of claims 1 to 4 wherein the amino-terminal tag sequence comprises a c-Myc epitope.

6. An expression vector according to any one of claims 1 to 5 wherein the amino-terminal tag sequence comprises the sequence as set forth in Seq ID No.3 or a fragment, mutant, variant or homologue thereof.

7. An expression vector according to any one of claims 1 to 6 wherein the signal sequence comprises the sequence as set forth in Seq ID No.5 or a fragment, mutant, variant or homologue thereof.

8. An expression vector according to any one of the preceding claims wherein the expression vector comprises a selectable marker.

9. An expression vector according to any one of the preceding claims wherein the expression vector is pIRES_STAR comprising the sequence as set forth in Seq ID No.2 or a fragment, mutant, variant or homologue thereof.

10. A host cell comprising the expression vector according to any one of claims 1 to 9.

11. A method for determining the utility of a nucleotide sequence of interest as a disease target, comprising the steps of:

(a) providing an expression vector according to any one of claims 1 to 9 comprising a nucleotide sequence of interest;

(b) transfecting a host cell with the expression vector;

(c) selecting for host cells expressing the nucleotide sequence of interest; and

(d) determining the expression profile of the protein expressed from the nucleotide sequence of interest.

12. A method for determining the utility of a nucleotide sequence of interest as a disease target, comprising the steps of:

(a) providing an expression vector according to any one of claims 1 to 9 comprising a nucleotide sequence of interest;

(b) transfecting a host cell with the expression vector;

(c) selecting for host cells expressing the nucleotide sequence of interest; and

(d) determining the expression profile of the protein expressed from the nucleotide sequence of interest; and

(e) purifying the protein expressed from the nucleotide sequence of interest.

13. A method according to claims 11 or 12 wherein the nucleotide sequence of interest is a disease target.

14. A method according to any one of claims 11 to 12 wherein the nucleotide sequence of interest is a tumour associated antigen.

15. A method according to any one of claims 11 to 14 wherein step (d) comprises FACS analysis of the binding of a labelled anti-amino-terminal tag antibody.

16. A method according to any one of claims 11 to 15 wherein step (d) comprises microscopic analysis of a labelled anti-amino-terminal tag antibody.

17. A method for preparing an antibody against a disease target polypeptide identified by the method of any one of claims 11 to 16 comprising the step of expressing the nucleotide sequence encoding the disease target polypeptide.

18. An antibody prepared by the method of claim 17.

19. An antibody according to claim 18 wherein the antibody is a monclonal antibody.

20. An antibody according to claim 19 wherein the antibody is a humanised antibody.

21. A kit for determining the utility of a nucleotide sequence of interest as a disease target comprising an expression vector according to any one of claims 1-9.

22. A library of nucleotide sequences of interest in an expression vector according to any one of claims 1 to 9.

23. A method of preparing a library according to claim 22 comprising the steps of:

(a) treating a nucleic acid sample to fragment the nucleic acid contained therein;

(b) ligating the nucleic acid fragments formed in step (a) in to a vector;

(c) transforming the vector into a host cell;

(d) selecting for host cells comprising the vector;

(e) identifying the signal sequence;

(f) amplifying the gene immediately downstream from the signal sequence;

(g) sub-cloning the gene in to an expression vector according to any one of claims 1-9; and

(h) determining the expression profile of the protein expressed from the nucleotide sequence of interest.

24. A disease target identified by the method of any one of claims 11 to 16 or 22.

25. An antibody according to any one of claims 18 to 20, wherein said antibody is linked to an agent.

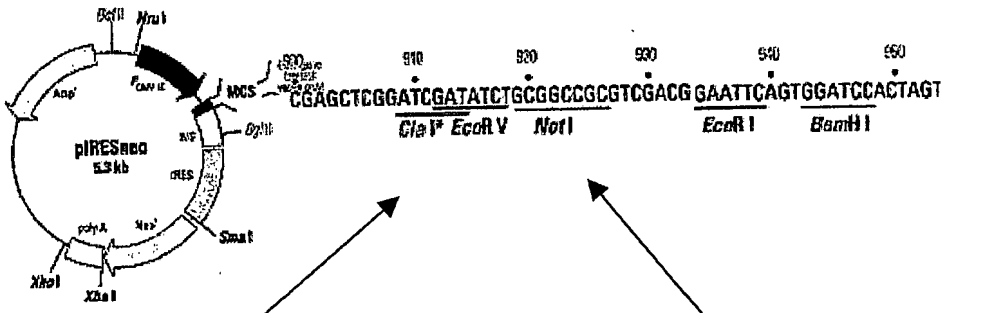
26. An antibody according to any one of claims 18 to 20 for use in the treatment of a disease.

27. Use of an antibody according to any one of claims 18 to 20 in the manufacture of a composition for the treatment of a disease.

28. A method of treating a disease in a subject comprising administering to said subject a medically effective amount of an antibody according to any one of claims 18 to 20.

29. An antibody according to claim 25, use of an antibody according to claim 26 or a method according to claim 27, wherein the disease is cancer.

Figure 1

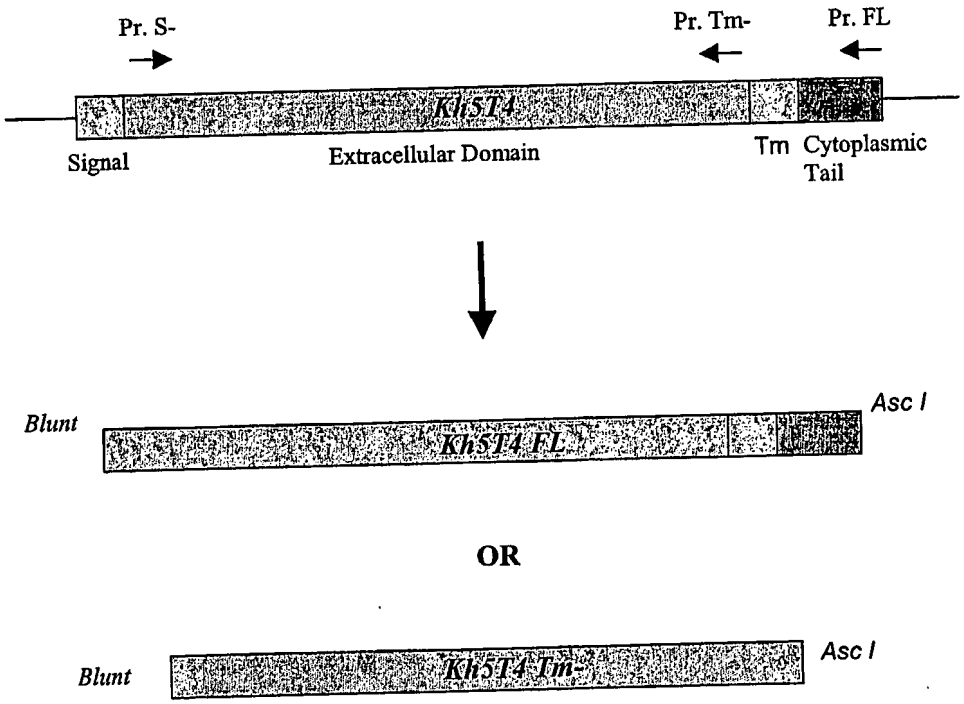



```
                M  G  W  S  C  I  I  L  F  L  V  A  T  A  T  G  V  H  S
1   CGATCCACCATGGGATGGAGCTGTATCATCCTCTTCTTGGTAGCAACAGCTACAGGTGTCCACTCC
      TAGGTGGTACCCTACCTCGACATAGTAGGAGAAGAACCATCGTTGTCGATGTCCACAGGTGAGG
    ClaI
```

6 His — Myc Epitope

```
                H  H  H  H  H  H  E  Q  K  L  I  S  E  E  D  L  A  F  K  R  D  I  G  A  P  R
67  CATCATCACCATCACCATGAACAAAAACTCATCTCAGAAGAGGATCTGGCGTTTAAACGCGATATCGGCGCGCCGC
    GTAGTAGTGGTAGTGGTACTTGTTTTTGAGTAGAGTCTTCTCCTAGACCGCAAATTTGCGCTATAGCCGCGCGGCGCCGG
                                                    PmeI    EcoRV   AscI    NotI
```

Figure 2



Pr. S: TCTTCTCCCACCTCCTCGG

Pr. FL: atatggcgcgccTCAGACATCCGAGTTAGAACT

Pr. Tm-: atatggcgcgccTCAGACATAAGAGGTTTGCAG

Figure 3

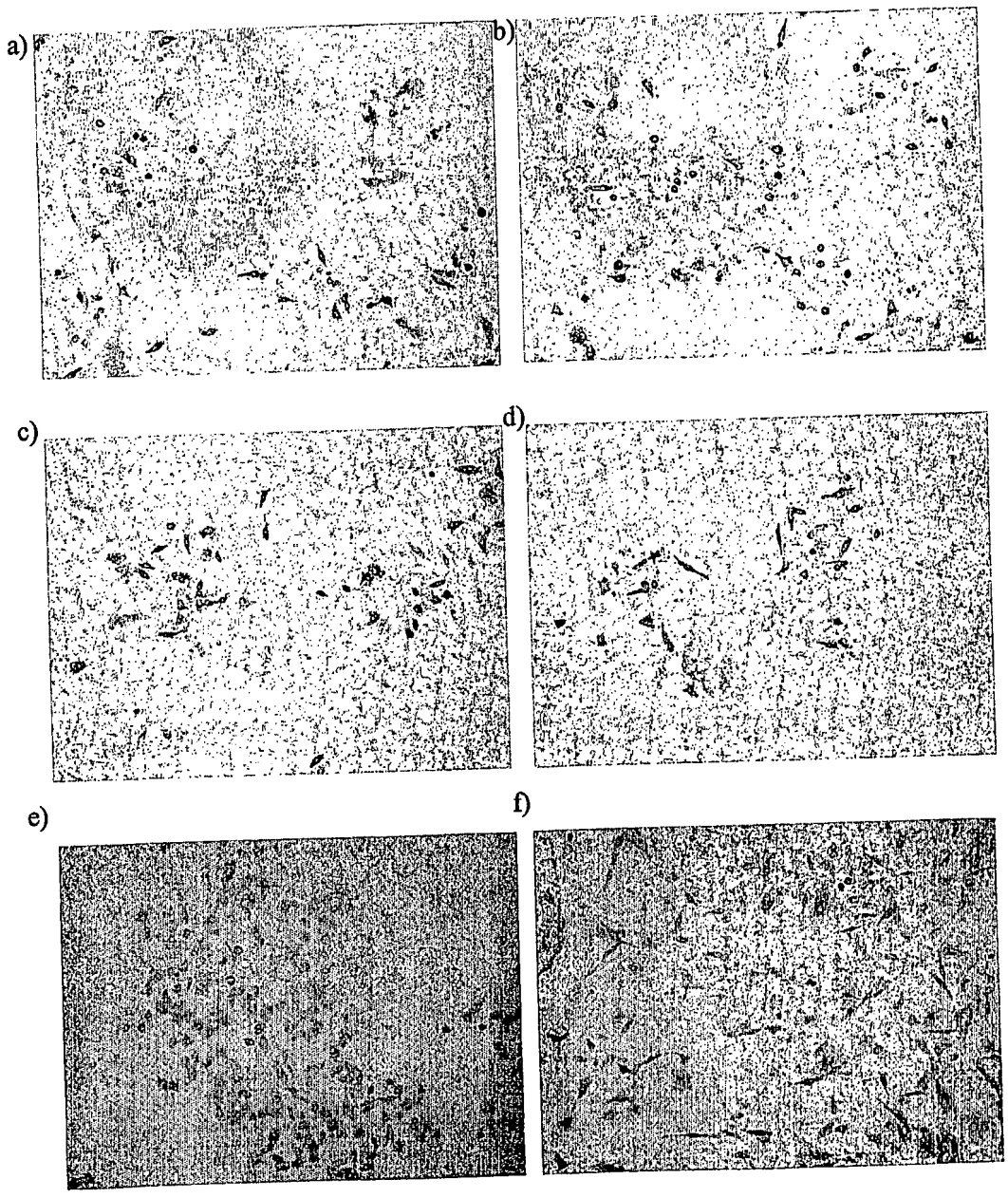


Figure 4

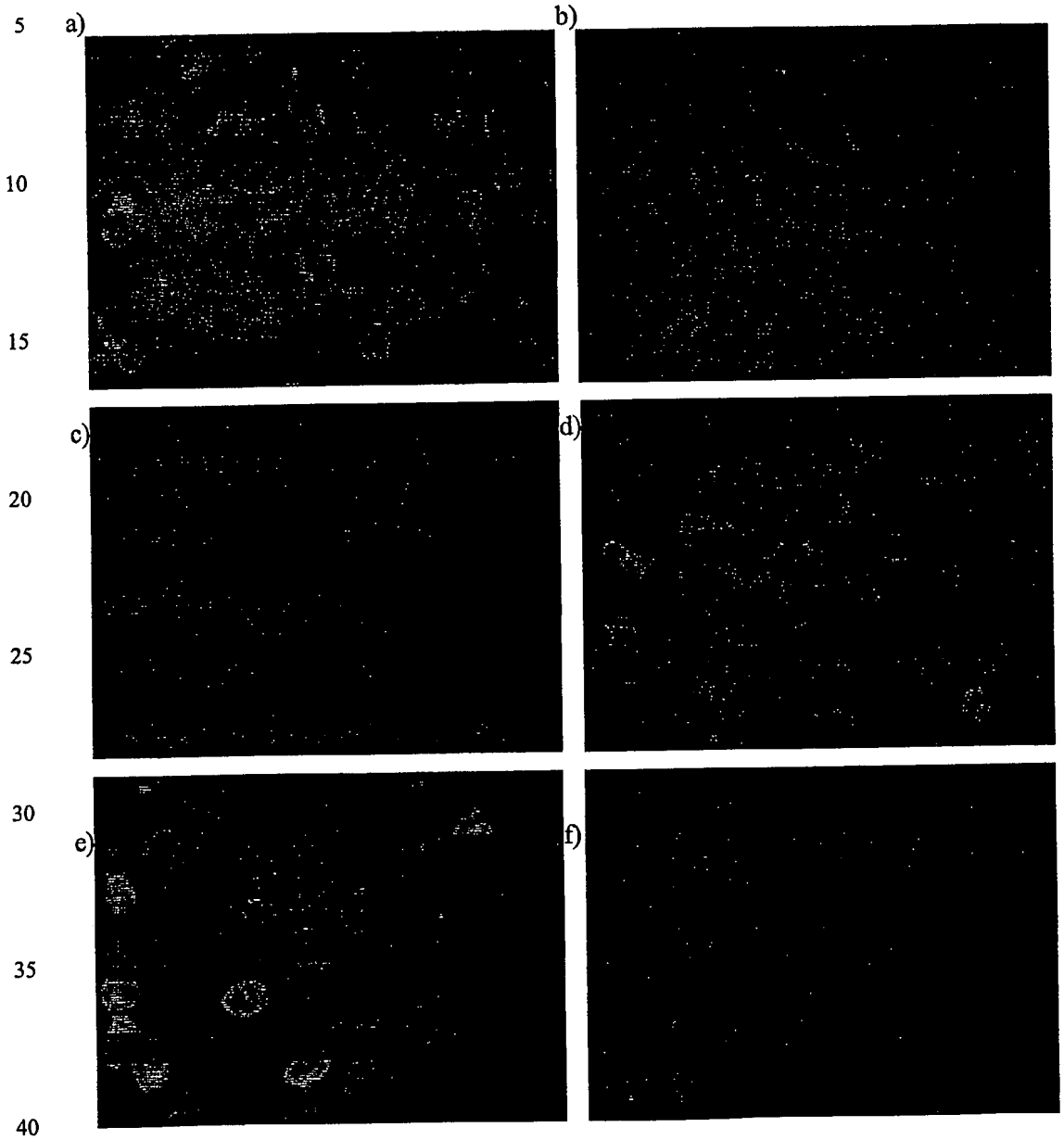


Figure 5A

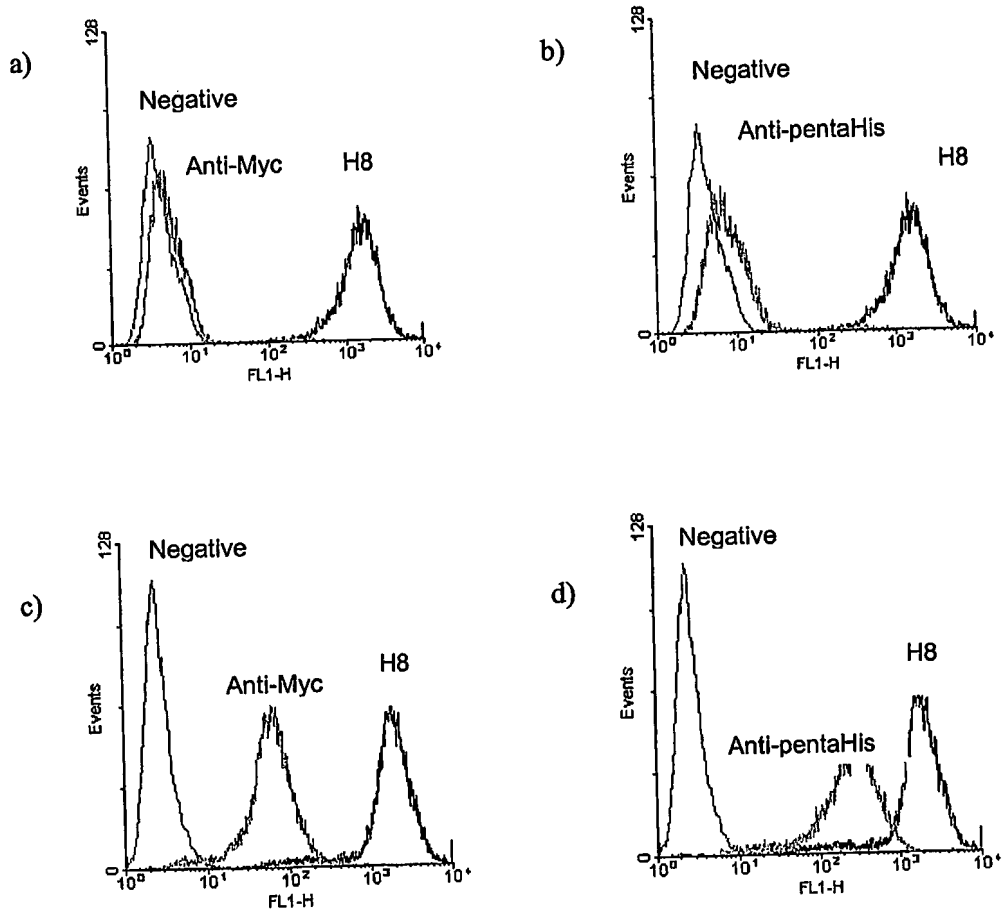
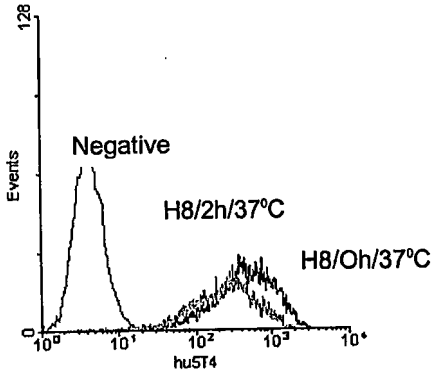

a)
Negative
Anti-Myc
H8
Events
FL1-H
b)
Negative
Anti-pentaHis
H8
Events
FL1-H
c)
Negative
Anti-Myc
H8
Events
FL1-H
d)
Negative
H8
Anti-pentaHis
Events
FL1-H

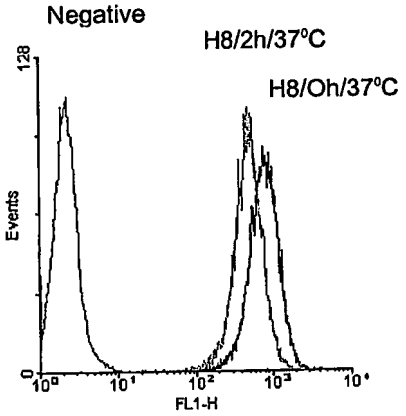
Figure 5B

6/7

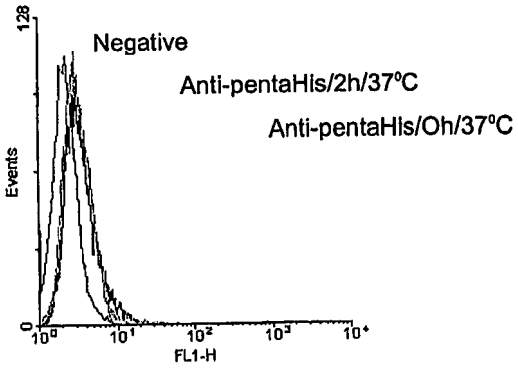
a)



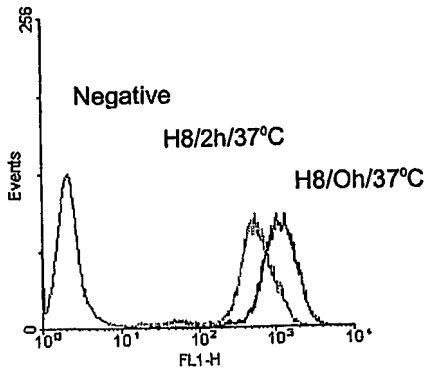
b)



c)



d)



e)

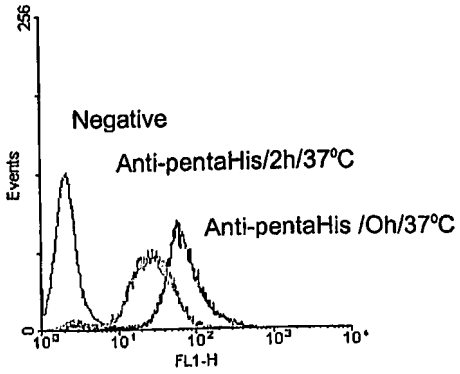
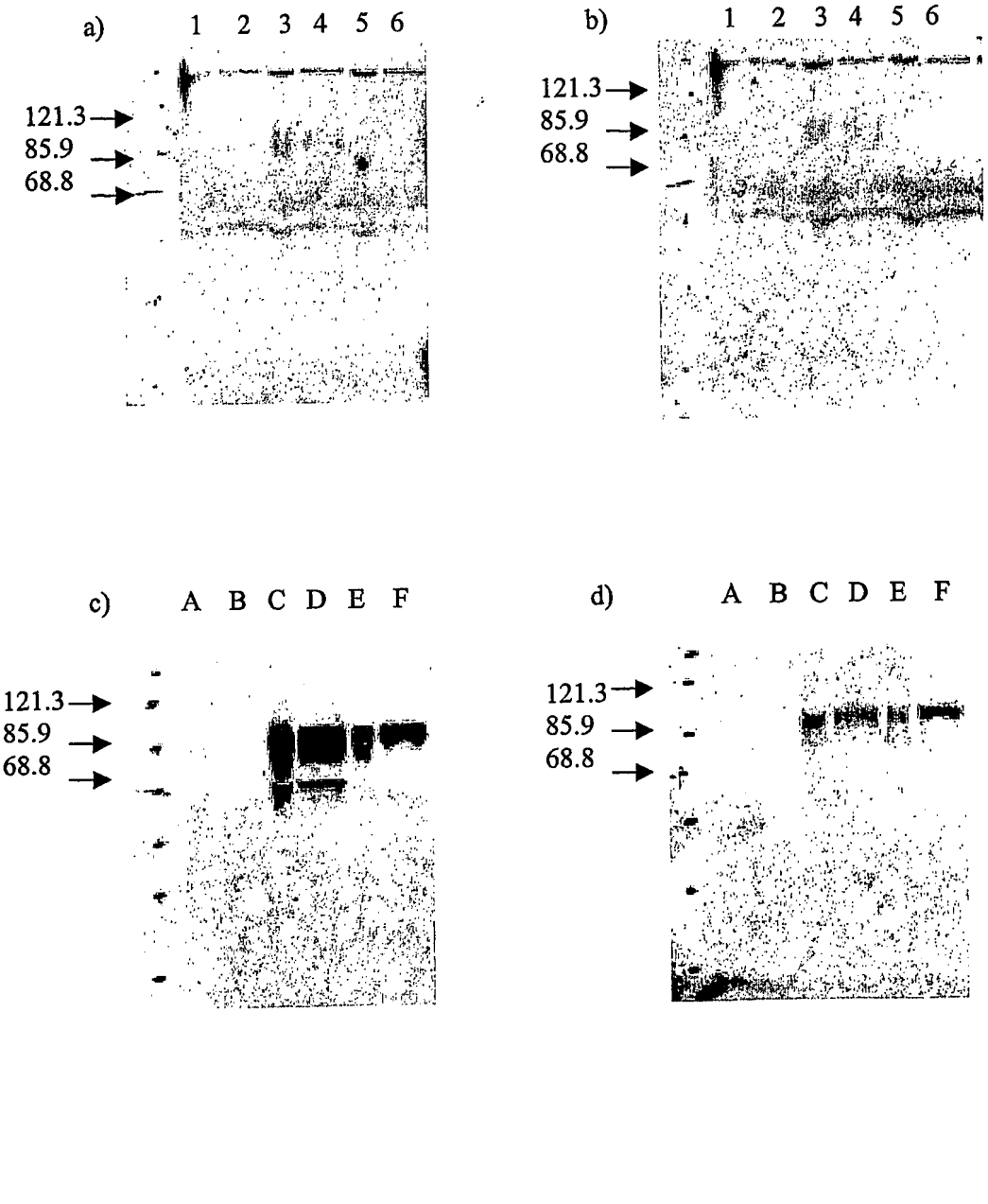


Figure 6


a)
1 2 3 4 5 6
121.3
85.9
68.8
b)
1 2 3 4 5 6
121.3
85.9
68.8
c)
A B C D E F
121.3
85.9
68.8
d)
A B C D E F
121.3
85.9
68.8

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 03/01659

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/85 C07K16/18 C12N15/62 C12N5/10 A61K39/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, MEDLINE, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 937 775 A (KYOWA HAKKO KOGYO KK ;BURNHAM INST (US)) 25 August 1999 (1999-08-25)	1-5,7,8, 10,21,23
Y	figures 7,9,11; examples 3,4 ---	6,9,22
X	HUSSY PETER ET AL: "Purification and in vitro-phosphalabeling of secretory envelope proteins E1 and E2 of hepatitis C virus expressed in insect cells." VIRUS RESEARCH, vol. 45, no. 1, 1996, pages 45-57, XP002252805 ISSN: 0168-1702 figure 1 ---	1,2,4,8, 10
	-/—	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 August 2003	17/09/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	ALCONADA RODRIG.., A

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 03/01659

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	SCHUSTER M ET AL: "Protein expression in yeast;comparison of two expression strategies regarding protein maturation." JOURNAL OF BIOTECHNOLOGY, vol. 84, no. 3, 2000, pages 237-248, XP002252806 ISSN: 0168-1656 page 1; table 1	11-16
Y	HOEDEMAEKER FLIP J ET AL: "A single chain Fv fragment of P-glycoprotein-specific monoclonal antibody C219. Design, expression, and crystal structure at 2.4 A resolution." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, no. 47, 21 November 1997 (1997-11-21), pages 29784-29789, XP002252807 ISSN: 0021-9258 page 29785, left-hand column, paragraph 3	6,9
Y	WO 00 14216 A (HADLOCK KENNETH G ;KECK ZHEN YONG (US); ZK PHARMACEUTICALS INC (US) 16 March 2000 (2000-03-16) page 2, line 15 -page 3, line 26	22
P,X	WO 02 61389 A (TANOX,INC.) 8 August 2002 (2002-08-08) page 4, paragraph 3 page 10, last paragraph page 8, paragraph 2 -page 9, paragraph 1 claims 7,8 figure 2	1,4,5,8, 10,11, 22,23

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17-20 and 24-29

Present claim 24 relates to a product defined by reference to a desirable characteristic or property, namely, that it can be identified by the method of any one of claims 11 to 16 or 22. The claims cover all products having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and disclosure within the meaning of Article 5 PCT for none of such products. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product/compound/method/apparatus by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search of the claimed scope impossible. Consequently, the search has not been carried out for this claim.

Present claims 17-20 and 25-29 relate to an antibody defined by reference to a desirable characteristic or property, namely, that it has been raised against a disease target identified by the method of any one of claims 11 to 16. The claims cover all antibodies having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and disclosure within the meaning of Article 5 PCT for none of such antibodies. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the antibody by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the claimed scope impossible. Consequently, the search has not been carried out for those claims.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB 03/01659

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 17-20 and 24-29
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 03/01659

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
EP 0937775	A	25-08-1999	EP	0937775 A2	25-08-1999
			JP	11236398 A	31-08-1999
			US	6420149 B1	16-07-2002
WO 0014216	A	16-03-2000	AU	5589399 A	27-03-2000
			CA	2341156 A1	16-03-2000
			EP	1109898 A1	27-06-2001
			WO	0014216 A1	16-03-2000
WO 0261389	A	08-08-2002	WO	02061389 A2	08-08-2002
			US	2003059834 A1	27-03-2003